Transactions of the ASME

PUBLISHING STAFF
Mng. Dir., Pub., J. J. FREY
Director, Technical Publishing
JOS. SANSONE
Managing Editor,
CORNELIA MONAHAN
Production Editor, SUZANNE MAGIDA
Prod. Asst., BETH DARCHI

APPLIED MECHANICS DIVISION

Chairman, WILLIAM G. GOTTENBERG
Secretary, CHARLES R. STEELE
Technical Editor, C. S. HSU
Associate Editors, J. D. ACHENBACH
S. T. ARIARATNAM
T. BELYTSCHKO
D. B. BOGY
R. J. CLIFTON
L. B. FREUND
MAURICE HOLT
W. D. IWAN
T. R. KANE
A. S. KOBAYASHI
S. LEIBOVICH
L. E. MALVERN
R. P. NORDGREN
R. H. PLAUT
R. T. SHIELD
C. T. SUN
J. S. WALKER
G. A. WEMPNER

BOARD ON COMMUNICATIONS
Chairman and Vice-President
M. J. RABINS

Members-at-Large
W. BEGELL
J. CALLAHAN
M. HELMICH
D. KOENIG
M. KUTZ
F. LANDIS
J. W. LOCKE
J. E. ORTLOFF
C. PHILLIPS
K. REID

President, SERGE GRATCH
Deputy/Exec. Dir.
PAUL ALLMENDINGER
Treasurer, ROBERT A. BENNETT

Journal of Applied Mechanics (ISSN 0021-8936) is edited and published quarterly at the offices of The American Society of Mechanical Engineers, United Engineering Center, 345 E. 47th St., New York, N.Y. 10017. ASME-TWX No. 710-581-5267, New York. Second Class postage paid at New York, N.Y., and at additional mailing offices.

CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.

PRICES: To members, $36.00, annually; to nonmembers, $72.00. Single copies, $24.00 each. Add $6.00 for postage to countries outside the United States and Canada.

STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1, Par. 3).

COPYRIGHT © 1982 by the American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the TRANSACTIONS OF THE ASME, JOURNAL OF APPLIED MECHANICS, and the author, and date of publication be stated.

INDEXED by the Engineering Index, Inc.

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 49 • NUMBER 4 • DECEMBER 1982

685 L. B. Freund Appointed Technical Editor of the Journal of Applied Mechanics

686 From the Technical Editor

TECHNICAL PAPERS

687 Similarity Analysis of Gas-Liquid Spray Systems
W. -S. Yeung

691 Strong Stability of Impulsively Initiated Couette Flow for Both Axisymmetric and Nonaxisymmetric Disturbances (82-WA/APM-20)
J. -C. Chen and G. P. Neitzel

697 A Model of the Turbulent Burst Phenomenon: Fully Developed Transitional Turbulent Flow Between Parallel Plates (82-WA/APM-10)
L. C. Thomas and H. M. Kadry

704 Experiment and Analysis of Instability of Tube Rows Subject to Liquid Crossflow (82-WA/APM-18)
S. S. Chen and J. A. Jendrzejczyk

710 Conservation Laws and Material Momentum in Thermoelasticity (82-WA/APM-26)
G. Francfort and A. Golebiewska Herrmann

715 The Stefan Problem With an Imperfect Thermal Contact at the Interface (82-WA/APM-17)
L. N. Tao

721 A Constitutive Model of Cyclic Plasticity With a Nonhardening Strain Region
N. Ohno

728 A Unified, Self-Consistent Theory for the Plastic-Creep Deformation of Metals
G. J. Weng

735 Optimal Cooling of Cross-Ply Composite Laminates and Adhesive Joints (82-WA/APM-24)
Y. Weitsman and B. D. Harper

740 Plasticity Analysis of Laminated Composite Plates (82-WA/APM-11)
Y. A. Bahei-El-Din and G. J. Dvorak

747 Plain-Strain Surface Waves in a Laminated Composite (82-WA/APM-21)
G. Herrmann and M. Hemami

754 Plastic Stress Intensity Factors for Small Scale Yielding in Antiplane Shear Caustics
P. S. Theocaris and C. I. Razem

761 Surface Instability and Splitting in Compressed Brittle Elastic Solids Containing Crack Arrays (82-WA/APM-16)
L. M. Keer, S. Nemat-Nasser, and A. Oranratnachai

768 Approximate Analytical Model of Anchor Pull-Out Test (82-WA/APM-28)
G. R. Miller and L. M. Keer

773 Effect of Shear and Rotary Inertia on Dynamic Fracture of a Beam or Plate in Pure Bending (82-WA/APM-9)
C. Levy and G. Herrmann

779 Extension and Torsion of Elastic Bars With Initial Twist
R. T. Shield

787 Stress Function Interface and Boundary Conditions in Anisotropic Materials (82-WA/APM-4)
E. E. Gdoutos and M. Kattis

792 Radiation From Expanding Circular Dislocation Loops and Elastic Precursor Decay (82-WA/APM-19)
X. Markenscoff and R. J. Clifton

797 Wavefront Analysis in the Nonseparable Elastodynamic Quarter-Plane Problems, I Part 1: The General Method (82-WA/APM-12)
J. Miklowitz

808 Wavefront Analysis in the Nonseparable Elastodynamic Quarter-Plane Problems, I Part 2: Wavefront Events in the Edge Uniform Pressure Problem (82-WA/APM-13)
J. Miklowitz

816 A Global Local Finite Element Analysis of Axisymmetric Scattering of Elastic Waves (82-WA/APM-3)
D. B. Goetschel, S. B. Dong, and R. Muki

821 Axisymmetric Elastic Waves Excited by a Point Source in a Plate (82-WA/APM-15)
R. L. Weaver and Y. H. Pao

837 Interferometric Determination of Curvatures of Flexed Plate (82-WA/APM-27)
Y. J. Chao, M. A. Sutton, and C. E. Taylor

843 Stability and Vibrations of Compressed, Aeolotropic, Composite Cylindrical Shells
J. B. Greenberg and Y. Stavsky

849 A Variable Parameter Incrementation Method for Dynamic Instability of Linear and Nonlinear Elastic Systems
S. L. Lau, Y. K. Cheung, and S. Y. Wu

854 Free Out-of-Plane Vibration of a Ring Elastically Supported at Several Points
T. Irie, G. Yamada, and H. Okada

861 Implicit Numerical Integration for Periodic Solutions of Autonomous Nonlinear Systems
G. A. Thurston

(Contents continued on Inside Back Cover)

Downloaded 04 May 2010 to 171.66.16.34. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

CONTENTS (CONTINUED)

867 **The Nature of the Temporal Solutions of Damped Distributed Systems With Classical Normal Modes** (82-WA/APM-14)
D. J. Inman and A. N. Andry, Jr.

871 **Compact Probabilistic Representation of Random Processes**
S. F. Masri and R. K. Miller

877 ***On-Line Control of Nonlinear Flexible Structures*** (82-WA/APM-25)
S. F. Masri, G. A. Bekey, and T. K. Caughy

885 **A Method of Analyzing Generalized Cell Mappings** (82-WA/APM-23)
C. S. Hsu, R. S. Guttalu, and W. H. Zhu

895 **A Probabilistic Theory of Nonlinear Dynamical Systems Based on the Cell State Space Concept** (82-WA/APM-22)
C. S. Hsu

903 **A Theory of Substructure Modal Synthesis** (82-WA/APM-29)
K. Kubomura

DESIGN DATA AND METHODS

910 **Natural Frequencies of Out-of-Plane Vibration Arcs**
T. Irie, G. Yamada, and K. Tanaka

BRIEF NOTES

A Simple and Effective Pipe Elbow Element—Pressure Stiffening Effects 914
K. J. Bathe and C. A. Almeida

Optimization of Reinforcement for a Class of Openings in Plate Structures 916
S. K. Dhir

Stability of the Vibrations of a Parameter-Excited System 920
M. Gürgoze

921 Axisymmetric Vortex Breakdown in Rotating Fluid Within a Container
H. J. Lugt and H. J. Haussling

923 Solution to the Backward-Kolmogorov Equation for a Non-stationary Oscillation Problem
G. P. Solomos and P. T. D. Spanos

925 A Note of the Length and Displacements of a Dugdale or Craze Zone
T. J. Lardner

DISCUSSIONS

927–930 Discussions on previously published papers by **S. K. Dhir, S. B. Dong** and **D. B. Goetschel, K. J. Bathe** and **C. A. Almeida,** and **T. C. T. Ting**

BOOK REVIEWS

931 ***Structural Modelling and Optimization*** by **D. G. Carmichael** . . . Reviewed by **J. E. Taylor**

Recent Developments in Thermomechanics of Solids edited by **G. Lebon** and **P. Perzyna** . . . Reviewed by **W. O. Williams**

932 ***Instabilities and Catastrophes in Sciences and Engineering*** by **J. M. T. Thompson** . . . Reviewed by **H. H. E. Leipholz**

Mechanics of Brittle Fracture by **G. P. Cherepanov** . . . Reviewed by **J. R. Rice**

Elementary Finite Element Method by **C. S. Desai** . . . Reviewed by **S. Kelsey**

933 ***Continuum Mechanics*** by **A. J. M. Spencer** . . . Reviewed by **R. L. Fosdick**

ERRATA

934 Erratum on "On the Logarithmic Singularity of Free-Edge Stress in Laminated Compoistes Under Uniform Extension," by **R. I. Zwiers, T. C. T. Ting,** and **R. L. Spilker,** and published in the September, 1982 issue
Erratum on "Analysis of the Weibull Distribution Function," by **K. T. Chang,** and published in the June, 1982 issue.

714, 727	Worldwide Mechanics Meeting List
703, 709, 778, 815, 848	Applied Mechanics Symposium Proceedings
860, 909, 913, 930, 934	Applied Mechanics Symposium Proceedings
767	Announcement – Standard International Units
842	Special Notice – Mandatory Excess Page Charge for Transactions
884	Change of Address Form

Downloaded 04 May 2010 to 171.66.16.34. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

L. B. Freund Appointed Technical Editor of the Journal of Applied Mechanics

Professor L. B. Freund has been appointed Technical Editor of the JOURNAL OF APPLIED MECHANICS effective January 1, 1983 for a five-year term. We are pleased that such a well-qualified person will be moving into this important position and at the same time, grateful to C. S. Hsu, the outgoing technical editor, for having done a magnificent job for the past six and one-half years.

Professor Hsu was appointed technical editor effective July 1, 1976 succeeding Phil Hodge, who himself served a five-year term. Readers and contributors to the Journal are fully aware of C. S. Hsu's success in maintaining the high standards and vitality of this important publication in applied mechanics. What may not be apparent is the level of patience and resourcefulness he has been required to sustain during these past few years of constantly inflating costs. Publication efforts of a society such as ASME that rely on a more or less fixed income are under extreme pressures in times of rising costs. We thank C. S. for bringing the Journal through this period and for agreeing to extend his original five-year commitment to the position.

In accordance with the rules of the Applied Mechanics Division, a committee was appointed in July 1981 to search for a new technical editor of the Journal. Phil Hodge agreed to serve as chairman of the committee whose members included Jan Achenbach, Michael Carroll, Tom Geers, and Jace Nunziato. This committee identified numerous qualified candidates who expressed a willingness to accept this important position if it were to be offered. Based on the search committee's recommendations, the Executive Committee selected and appointed Prof. Freund as Technical Editor. This position also requires a commitment from the editor's home institution and we are pleased that Brown University has agreed to provide the editorial office for our journal.

Ben Freund is Professor of Engineering and Chairman of the Executive Committee-Division of Engineering at Brown, having been affiliated with that institution since 1967. Freund's initial education in mechanics was at the University of Illinois and he received his Ph.D in applied Mechanics from Northwestern University in 1967. He is a Fellow of ASME and has served as Associate Editor of the JOURNAL OF APPLIED MECHANICS since 1979.



L. B. Freund

Professor Freund has published extensively on wave propagation and crack propagation in elastic solids. He has discussed applications of these subjects in geophysics and has done analysis of running fractures in gas pressurized pipes. More recently his research has concentrated on the propagation of waves in elastic-plastic materials as well as dynamic crack growth in such materials.

We consider publication of the Journal to be one of the most important activities of the Division and sincerely believe that its leadership is in good hands for another five years, at least, with Ben Freund's appointment as technical editor.

Bill Gottenberg
for the Executive Committee

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

From the Technical Editor

As announced on the preceding page, Professor L. B. Freund has been appointed as the new Technical Editor of this Journal. He will assume the editorship on January 1, 1983. Starting from that date, all manuscripts and correspondence with the Office of the Technical Editor should be sent to

Professor L. B. Freund, Technical Editor
Journal of Applied Mechanics
Division of Engineering
Brown University
Providence, RI 02912

The current plan is to physically move the files of the Office from Berkeley, California to Providence, Rhode Island around December 20, 1982. For this reason it will help greatly if the authors could refrain from sending material to the Technical Editor between December 15 and December 31.

Looking back now at the last six and one-half years of my term as the Technical Editor of the Journal, I find the work to be one of the most rewarding and interesting experiences, and I shall always feel privileged to have had the opportunity to serve the Journal in this capacity. The editorship does involve a great deal of work, but my work was considerably eased because of the remarkably sensible operation system Phil Hodge had set up for the Journal when he was the Technical Editor. The longer I work on the job, the more I appreciate his wisdom and foresight in devising the system. During my tenure I have followed very much the mode of operation he had initiated, with only very minor changes dictated by the new policies of the Society. For this and for his constant willingness to give me advice and help when needed, I want to express to him my gratitude.

With regard to the operations of the Journal, a few words about our editorial system might be in order here. While all the authors know this system well, many of our readers may not. The JOURNAL OF APPLIED MECHANICS follows a strong associate editor system. Each paper is assigned to an Associate Editor who is either an expert himself or a highly knowledgeable person in the subject area of the paper. In our system it is the Associate Editor who carries the burden and the responsibility of selecting the reviewers, evaluating the reviews, and making the final editorial decision. In the last several years it has been my good fortune to work with an outstanding group of Associate Editors. Those who have finished their terms are:

D. D. Joseph, University of Minnesota
Eli Reshotko, Case Western Reserve
James Rice, Brown; now at Harvard
J. Dundurs, Northwestern University
Don Griffin, Pittsburgh, Pa.
Thomas Lardner, University of Massachusetts
Stephen Cowin, Tulane University
J. W. Dally, University of Maryland; now Washington, D.C.
T. K. Caughey, California Institute of Technology
J. W. Hutchinson, Harvard University
P. A. Libby, University of California at San Diego
J. J. McCoy, The Catholic University of America
T. C. T. Ting, University of Illinois at Chicago Circle
M. S. Plesset, California Institute of Technology
Robert Plunkett, University of Minnesota
Owen Richmond, U.S. Steel Research Laboratory
C. R. Steele, Stanford University
U.S. Lindholm, Southwest Research Institute

Those who are on the current Editorial Board are:

S. T. Ariaratnam, University of Waterloo
Ted Belytschko, Northwestern University
D. B. Bogy, University of California, Berkeley
R. J. Clifton, Brown University
Maurice Holt, University of California, Berkeley
W. D. Iwan, California Institute of Technology
T. R. Kane, Stanford University
A. S. Kobayashi, University of Washington
S. Leibovich, Cornell University
L. E. Malvern, University of Florida
R. P. Nordgren, Shell Development Company
R. H. Plaut, VPI & State University
R. T. Shield, University of Illinois at Urbana-Champaign
C. T. Sun, Purdue University
J. S. Walker, University of Illinois at Urbana-Champaign
G. A. Wempner, Georgia Institute of Technology

To all these past and current Associate Editors, thank you for your hard work and professional dedication.

Another group of unsung heroes of the operation of the Journal is, of course, the large corps of reviewers the editors are able to call on to help. I believe an excellent tradition has been established for the Journal, so the reviewers all take their responsibility most seriously to maintain the high standard of the Journal, and to be fair. In my view this is truly the strength of the Journal. The readers probably have seen in the issues some of the authors' acknowledgments of very helpful suggestions from the reviewers which have materially improved the papers. To all these reviewers who service is gratefully acknowledged each year in the September issue, thank you again for your effort and time.

While the Associate Editors and the Technical Editor have to decide which manuscripts to accept and which to decline, once a paper is accepted, the long and arduous copy editing and printing process is shifted to the New York Editorial Office. In this respect, the JOURNAL OF APPLIED MECHANICS has been blessed with excellent production editors. Ronnie Heaney was in that position for 13 years until she left ASME in 1981 after 26 years of service. Ronnie was succeeded by Suzanne Magida, our current Production Editor. Both have given the Journal excellent service, and I want to take this opportunity to thank them both on behalf of the Journal and the Applied Mechanics Division.

Closer to home, I want to express my warm appreciation to my secretary and assistant, Janie Ellison. She started with me in 1976 to set up the Technical Editor's Office in Berkeley. Besides overseeing the smooth flow of the manuscripts from the authors, to and from the associate editors, and finally to the New York Office, she also attends to a myriad of other details of the office. On several occasions when I was away from Berkeley on extended trips, she ran the operations of the Office by herself. For all these years of long and efficient help, my heartfelt thanks.

As mentioned by Bill Gottenberg, the Technical Editor must have the support of his home institution in order to do his job. On this, I want to thank Professors C. L. Tien and I. Finnie, the past and current chairmen of the Department of Mechanical Engineering at Berkeley, for their strong support. This is particularly appreciated in this era of stringent budget situations everywhere. At the Society level I want to thank the Executive Committee of the Applied Mechanics Division for their trust in appointing me to this editorship and for their continual support throughout these years.

Finally, I send Ben Freund my best wishes for the future. Ben is no stranger to this Journal. He was one of the very hardworking reviewers before he took on the associate editor assignment in 1979. I join with many others in wishing him success in the coming years.

C. S. Hsu
Technical Editor

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Similarity Analysis of Gas-Liquid Spray Systems[1]

W.-S. Yeung[2]
Department of Mechanical Engineering,
University of Lowell,
Lowell, Mass. 01854
Assoc. Mem. ASME

This paper discusses the fluid mechanics of the droplet-dispersed region of a liquid-spray system. Similarity solutions are sought under various assumptions for two types of common spray: the conical spray and the flat spray. Following the classical approach of Schlichting, for turbulent jet analysis, we obtain a similarity solution for the flat spray under the condition that the droplet motion is not influenced by gravity and interfacial drag. The solution is quite realistic from a physical standpoint. Finally, comparison between the theory and limited experimental data was found to give good agreement.

Introduction

This paper studies the fluid mechanics of a gas medium dispersed with liquid droplets that are formed when a liquid is sprayed through a pressure-atomizing nozzle. Briefly, the liquid first emerges from the nozzle in the form of a jet or a sheath which quickly disintegrates into droplets of various sizes due to instability effects. The region where the liquid jet or sheath has not yet disintegrated is called the near region and the region where the liquid exists as droplets is called the far region. A schematic diagram of a liquid spray is shown in Fig. 1. We shall only consider the dynamics of the far region.

The droplets possess initial velocity when they are formed from the disintegration process. As they travel downstream, they drag along the ambient gas and consequently set up an overall induced-gas flow field. This induced-gas motion in turn modifies the droplet trajectories. The subsequent fluid mechanics of the far region is then governed by the interaction between the droplet and gas phases.

Numerous analytical and experimental studies have been made on the aerodynamic behavior of liquid sprays. Among these are the work of Rothe and Block [1] based on a one-dimensional model; Boysan and Binark [2], and Alpert and Mathews [3] based on a three-dimensional numerical solution. Experimentally, Heskestad, Kung, and Todtenkopf [4] measured the air velocity inside a water spray. Measurements of the induced air motion were previously obtained by Rashbash and Stark [5] who also found that the induced air motion is dynamically quite similar to that of a turbulent single-phase gas jet.

In the present study, we seek similarity solutions for the far region under various physical situations and assumptions. The analysis follows basically the approach of Schlichting [6] for turbulent jet analysis.

Theory

We consider two types of liquid spray: the conical spray and the flat spray. The physical system is shown in Fig. 1. A nozzle discharges liquid vertically downward into a semi-infinite region of incompressible, ambient gas. The volume occupied by the liquid droplets inside the spray envelope is assumed negligible. Furthermore, the flow field is assumed to be two-dimensional for the flat spray and axisymmetric for the conical spray.

Denote x as the distance vertically downward from the nozzle and y as the normal distance from the spray axis. The velocity components of either phase are denoted by u and v in the directions of increasing x and y, respectively. In view of the experimental findings of [5] mentioned in the In-

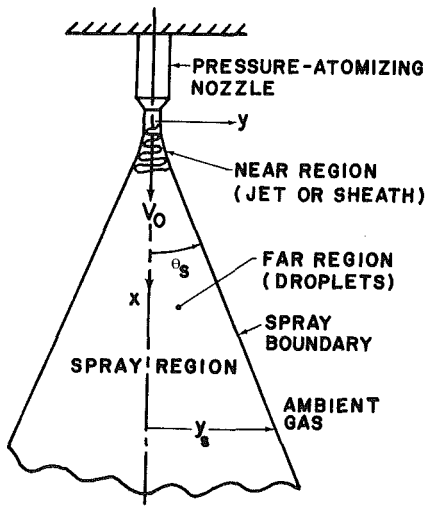


Fig. 1 Notations of a liquid spray

[1]This work was performed at Thayer School of Engineering, Dartmouth College, Hanover, N.H. 03755, under a research grant from the Electric Power Research Institute, Palo Alto, Calif. with Dr. K. H. Sun as Program Director.

[2]Formerly, Research Assistant Professor of Engineering, Dartmouth College, Hanover, N.H. 03755.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 1981; final revision, April, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

troduction, the general governing equations for the gas phase are taken to be those describing a single-phase turbulent jet, taking into account the interfacial drag force:

$$\partial(uy^\sigma)/\partial x+\partial(vy^\sigma)/\partial y=0 \tag{1}$$

$$u(\partial u/\partial x)+v(\partial u/\partial y)=(1/y^\sigma)\partial(y^\sigma \nu_{\text{eff}}\partial u/\partial y)/\partial y - (\rho_d/\rho)(u-u_d)/\tau_m \tag{2}$$

The last term in equation (2) is the interfacial drag force and is not present in the single-phase, turbulent-jet equation. The droplets are treated as rigid spheres with a constant mean diameter. Thus, formulations written for a gas-particle mixture can be used here. We use the somewhat simplified and common formulation, similar to that used by Singleton [7], among others. The droplet phase is treated as a continuum, and the governing equations are:

$$\partial(\rho_d u_d y^\sigma)/\partial x+\partial(\rho_d v_d y^\sigma)/\partial y=0 \tag{3}$$

$$u_d(\partial u_d/\partial x)+v_d(\partial u_d/\partial y)=(u-u_d)/\tau_m \tag{4}$$

$$u_d(\partial v_d/\partial x)+v_d(\partial v_d/\partial y)=(v-v_d)/\tau_m \tag{5}$$

In equations (1)–(5), ρ is the phase density, ν_{eff} is the turbulent viscosity, and the subscript d refers to the droplet phase. The unsubscripted variables refer to the gas phase, unless otherwise stated. The geometrical factor σ takes the value zero for flat sprays and unity for conical sprays. In the foregoing, gravity effects are neglected. Its implications will be discussed later. In equations (2), (4), and (5), τ_m is called the momentum equilibration time and is given by

$$\tau_m=(\rho_f D_d^2/18\mu)(24/|\text{Re}|C_D) \tag{6}$$

where C_D is an appropriate drag coefficient, taken as [8]

$$C_D=24/|\text{Re}|(1+0.15|\text{Re}|^{0.687}), \quad |\text{Re}|<1000 \tag{7}$$

$$C_D=0.44, \quad |\text{Re}|>1000, \tag{8}$$

Re is the Reynolds number, defined as

$$\text{Re}=\rho(u-u_d)D_d/\mu, \tag{9}$$

μ is the gas dynamic viscosity, ρ_f is the liquid density, and D_d is a mean droplet diameter. Equations (1)–(5) are the basic governing equations inside the spray envelope. Outside the spray envelope, we assume no droplets are present. Thus, equations (3)–(5) are no longer needed, and ρ_d is identically zero in equation (2). Equation (1) remains unchanged. The boundary conditions are

$$u=0, \quad y\to\infty \tag{10}$$

and

$$\partial u/\partial y=\partial u_d/\partial y=\partial\rho_d/\partial y=v=v_d=0, \quad y=0 \tag{11}$$

due to symmetry.

Similarity Analysis

We begin by defining a stream function, ψ, such that

$$u=1/y^\sigma(\partial\psi/\partial y), \quad v=-1/y^\sigma(\partial\psi/\partial x) \tag{12}$$

We seek solutions of the following types:

$$\psi\sim x^p f(\eta) \tag{13}$$

$$\rho_d\sim x^q G(\eta) \tag{14}$$

where η is the similarity variable, defined as

$$\eta=y/x^n \tag{15}$$

Following Schlichting, we express ν_{eff} as

$$\nu_{\text{eff}}=\kappa x^m \tag{16}$$

where κ is an empirical coefficient. The main objective is to reduce equations (1)–(5) to a system of ordinary differential equations in η by assigning appropriate values to the x exponents p, q, m, and n. Also note that u_d and v_d must have the same x dependence as u and v, respectively, due to the relative velocity term in equations (2), (4), and (5).

The appropriate values of the x exponents can be derived as follows. From equation (2), the balance between the inertial and the viscous forces gives

$$p-n(\sigma-1)-m=1 \tag{17}$$

and the balance between the inertial and the drag forces gives

$$p-n(\sigma+1)-q=1 \tag{18}$$

Equations (4) and (5) lead to a single relation

$$p-n(\sigma+1)=1 \tag{19}$$

Since there are no droplets outside the spray envelope, ρ_d, therefore, $G(\eta)$ in equation (14) must equal zero there. Thus, denote η_s as the value of η at the spray boundary and write $G(\eta)$ as

$$G(\eta)=\Gamma(\eta-\eta_s)\bar{G}(\eta) \tag{20}$$

where $\Gamma(\eta)$ is unity for $\eta<0$ and zero for $\eta>0$. If G is truly a function of η alone, η_s must be a constant, denoted by C. It follows from equation (15) that the spray boundary is given by

$$y_s=Cx^n \tag{21}$$

Equation (21) fixes the possible shape of the spray boundary for similarity solutions. Finally, from mass conservation of liquid, we have,

$$2J\int_0^{y_s}\rho_d u_d y^\sigma dy=\dot{W}_L \quad \text{for all} \quad x \tag{22}$$

where $J=\pi$ for conical sprays and $J=\delta$, with δ being the thickness of the flat spray, for flat sprays. $\dot{W}_L$ is the liquid mass flow rate through the nozzle. Changing the variable of integration to η and using (14) and (15), we arrive at the following result:

$$p+q=0 \tag{23}$$

Equations (17)–(19), (21), and (23) are the basic relations for the similarity analysis. In deriving these equations, κ and τ_m have been assumed constant. τ_m is constant only when the Reynolds number given by equation (9) approaches zero, i.e., the Stokes flow regime. However, as we shall see, the analysis is not restricted to the Stokes flow regime as long as τ_m does not depend on x explicitly, or τ_m is a function of η alone. Constant κ implies that the turbulence level is uniform across the whole section at a given distance x. However, in a spray system, the gas motion appears to be much more turbulent inside the spray region than outside the spray region [2]. To be more realistic, κ should be taken as a decreasing function of η. We shall return to these points later.

Similarity Solution

Two different cases have been considered. The first corresponds to the situation where the drag force influences the droplet motion. However, it has been shown that no compatible similarity solutions of the types given by (13) and (14) are possible. The second case corresponds to the physical situation where the droplet equilibration time is very large ($\tau_m\to\infty$). The droplet motion is then entirely inertia dominated and is not influenced by the drag force over a sufficiently large distance. With the gravity force also neglected in this analysis, we conclude that the droplet velocity remains constant over this inertia dominated region. Moreover, the droplet trajectories are straight lines. In this case, a compatible similarity solution was found for the flat spray. We shall proceed to discuss this solution. The reader is referred to Yeung [9] for details of other solutions that are deemed unrealistic. The Appendix of this paper contains an example of an unrealistic similarity solution for the conical spray.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

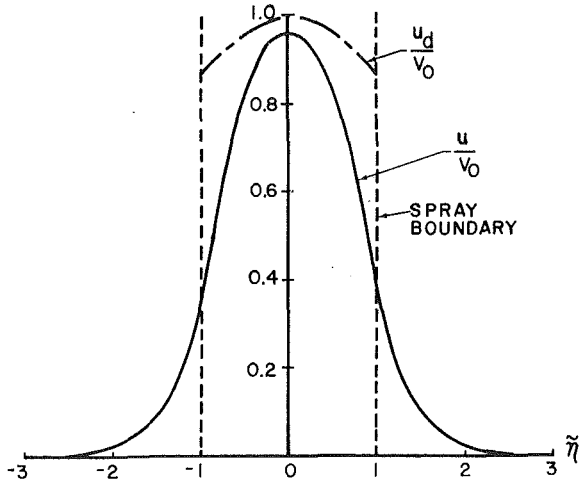


Fig. 2 Similarity velocity profile for the plane turbulent spray. $\kappa/V_0 = 0.1$, $Re_D = 1000$, $\kappa = 1.0$, and $\theta_s = 0.524$ rad.

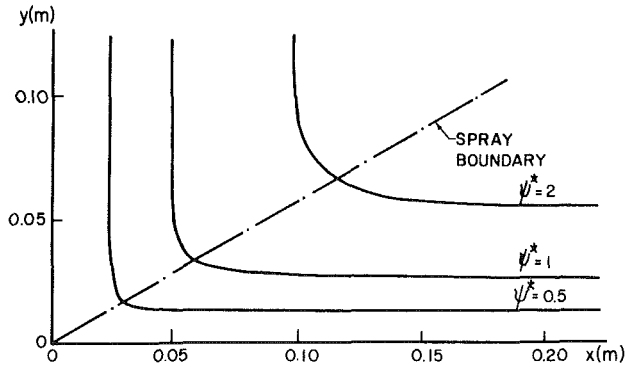


Fig. 3 Gas streamlines for the plane turbulent spray of Fig. 2. $\psi^* \equiv \psi/V_0$.

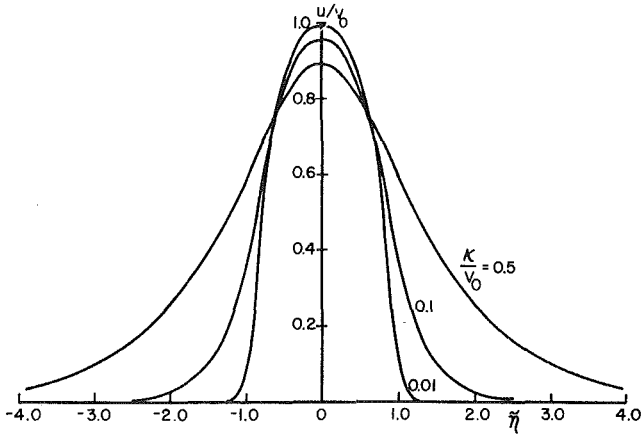


Fig. 4 Effect of the turbulent viscosity on the axial gas velocity

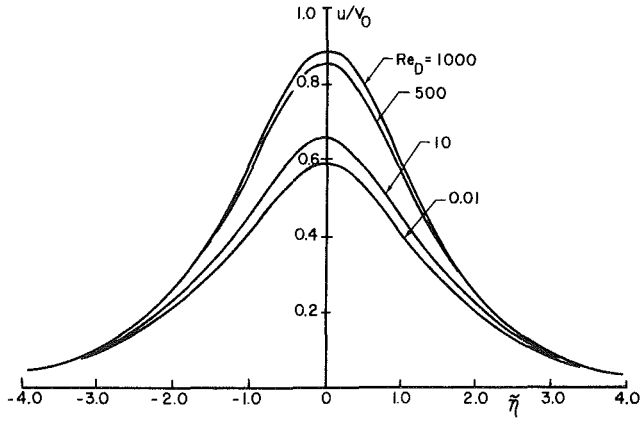


Fig. 5 Effect of Re_D on the axial gas velocity

Solution to Flat Turbulent Spray When $\tau_m \to \infty$

Since equation (19) is the result of balancing the particle inertial force with the drag force, it is not applicable for $\tau_m \to \infty$. However, in this case, the droplet trajectories are straight lines. Hence $n = 1$ from equation (21). Equations (17), (18), and (23) then give $p = 1$, $q = -1$, and $m = 1$. The similarity variable is then given by y/x. The droplet velocity components are simply given by

$$u_d = V_0/\sqrt{1+\eta^2} \tag{24}$$

$$v_d = V_0\eta/\sqrt{1+\eta^2} \tag{25}$$

where V_0 is the droplet speed at the breakup point, as shown in Fig. 1. The droplet phase density can now be written, without loss of generality, as

$$\rho_d = (\rho D_s/x)\,G(\eta) \tag{26}$$

where D_s is an equivalent diameter of the nozzle, usually supplied by the nozzle manufacturer. Equation (26) satisfies the droplet continuity equation (3) regardless of the form of $G(\eta)$. In practice, one should choose a functional form of $G(\eta)$ or $\bar{G}(\eta)$ which approximates physical observations. Experiments with solid-cone, air-water sprays typically show that the vertical droplet mass flux at a given distance is normally distributed with y [10]. Noting that the vertical mass flux is given by $\rho_d u_d$, $\bar{G}(\eta)$ can be deduced as

$$\bar{G}(\eta) = \frac{\lambda \dot{W}_L}{2\rho D_s \delta V_0 \tan\theta_s}\sqrt{1+\eta^2}\,e^{-\chi\eta^2/\tan^2\theta_S} \tag{27}$$

where

$$\lambda = \{\tfrac{1}{2}\sqrt{\pi/\chi}\,\mathrm{erf}(\sqrt{\chi})\}^{-1} \tag{28}$$

and erf denotes the usual normalized error function. χ is an empirical constant in the foregoing and θ_s is the spray half angle. Since the mass flux is virtually zero at the spray boundary, χ is set equal to 3 in this analysis.

We now consider the induced gas motion. If we rewrite the stream function as $V_0 x f(\eta)$, evaluate the corresponding gas velocity components, and use the various expressions for ν_{eff}, ρ_d, and u_d, we can reduce equation (2) to the following ordinary differential equation:

$$-\tilde{f}(\tilde{\eta})\tilde{f}''(\tilde{\eta}) = \kappa/V_0(\tilde{f}'''(\tilde{\eta})/\tan^2\theta_s) - \lambda\Gamma(\tilde{\eta}-1)\omega e^{-\chi\eta^2}\cdot\{\sqrt{1+\tilde{\eta}^2\tan^2\theta_s}\,\tilde{f}(\tilde{\eta}) - 1\} \tag{29}$$

where

$$\tilde{f}(\tilde{\eta}) = f(\eta)/\tan\theta_s, \quad \tilde{\eta} = \eta/\tan\theta_s \tag{30}$$

and

$$\omega = (\dot{W}_L/2\rho D_s \delta V_0 \tan^2\theta_s)(D_s/V_0\tau_m) \tag{31}$$

As mentioned before, τ_m can be a function of η. It can be easily shown that the two phasic velocities are functions of η alone, so that the relative velocity $(u-u_d)$ or $(v-v_d)$ is also a function of η alone. Thus, from equations (6)-(9), τ_m is indeed a function of η and equation (29) is still valid. Hence, the present solution is not limited to the Stokes flow regime. This greatly improves the practicality of the solution. The boundary conditions are, from equations (10) and (11),

$$\tilde{f}(0) = \tilde{f}''(0) = 0; \quad \tilde{f}'(\infty) = 0 \tag{32}$$

For numerical purpose, it is convenient to rewrite equation (31) as

$$\omega = \xi(|\mathrm{Re}|C_D)/24 \tag{32}$$

where ξ is a system constant

$$\xi = 9\mu\dot{W}_L/\rho\rho_f\delta V_0^2 D_d^2 \tan\theta_s \tag{33}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

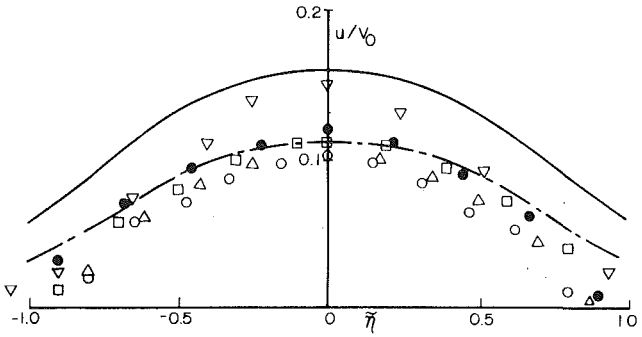


Fig. 6 Comparison of the similarity solution with experimental results [11]. Experimental data: ▽, x = 0.91m; •, 1.07 m; □, 1.22m; △, 1.37m; ○, 1.52m; theoretical:——, $D_d = D_s/10$, $\Omega_1 = 1$, $\Omega_2 = 20$; — — —, $D_d = D_s/5, \Omega_1 = 0.5, \Omega_2 = 10$.

by using equation (6). The relative Reynolds number can also be written as

$$\mathrm{Re} = \mathrm{Re}_D (f' - 1/\sqrt{1+\eta^2}) \tag{34}$$

where Re_D is given by

$$\mathrm{Re}_D = \rho V_0 D_d / \mu \tag{35}$$

Given values of θ_s, ξ, Re_D, and the appropriate form of κ/V_0, equation (29) can be integrated numerically. Solutions have been obtained using the traditional shooting method.

Results and Discussions

κ was first assumed to be constant. Figure 2 shows a typical case of the droplet and gas velocity profiles for the particular set of values of θ_s, κ/V_0, ξ, and Re_D shown. The bell-shape profile of u is in general qualitative agreement with experimental observations of spray systems. The corresponding gas streamlines in the physical plane are shown in Fig. 3. The gas approaches the spray almost perpendicularly to the spray boundary and then turns abruptly to a direction parallel to the spray axis. Figure 4 shows the velocity profiles at different values of κ/V_0. The effect of Re_D on the gas profiles is shown in Fig. 5.

The appropriate form of κ must be eventually estimated by comparison with experimental data, as was done in turbulent single-phase jets. An experimental study of the flat turbulent spray was carried out by Chuah et al. [11] using several "Unijet" nozzles manufactured by the Spraying Systems Co., Nashua, N. H. The water volumetric flow rates from the nozzles were less than 1.26×10^{-4} m^3/s typically. The existence of similarity solutions was not evident for all the nozzles tested. Nevertheless, comparisons were made with those results that indicate the plausibility of similarity solutions. Since the turbulence level inside the spray envelope is higher than that outside, κ is arbitrarily taken as

$$\kappa/V_0 = \Omega_1/1 + \Omega_2 \tilde{\eta} \tag{36}$$

where Ω_1 and Ω_2 are adjustable constants to fit experimental data. Figure 6 shows a typical comparison between the analytical and experimental results, with $\Omega_1 = 1$ and $\Omega_2 = 20$. The similarity solution overpredicts the experimental results. Many combinations of Ω_1 and Ω_2 have been tried, but with little success in matching the overall velocity profile. A possible explanation for this overprediction is the incorrect estimation of the parameters ξ and Re_D for the nozzles used in [11]. A particular source of uncertainty is the droplet diameter D_d which directly affects the values of ξ and Re_D. In the foregoing, D_d was taken as one-tenth of D_s. By taking D_d as one-fifth of D_s, the similarity solution was obtained for the case of $\Omega_1 = 0.5$ and $\Omega_2 = 10$. The result is also shown in Fig. 6 and is seen to be in fair agreement with the experimental data.

Conclusions

In summary, a similarity solution was obtained for the flat turbulent spray. The existence of this solution and its validity are at least verified by the limited available experimental results. There is a degree of empiricism associated with the present analysis concerning the appropriate form of the turbulent viscosity. A better understanding of two-phase flow turbulence is needed. Another difficulty associated with the present study is the uncertainty of specifying certain system parameters, such as the mean droplet size and the initial droplet velocity.

Although these drawbacks exist, the present analysis should be of theoretical interest. Similarity solutions are scarce in fluid mechanics and even more so in the field of two-phase flow. The analytical technique presented here is both simple and systematic, and the solution found is quite realistic from a physical standpoint.

Acknowledgment

Special thanks are due to Professor G. B. Wallis for directing the author to this problem.

References

1 Rothe, P. H., and Block, J. A., "Aerodynamic Behavior of Liquid Sprays," *International Journal of Multiphase Flow*, Vol. 3, 1977, pp. 263–272.

2 Boysan, F., and Binark, H., "Prediction of Induced Air Flows in Hollow Cone Sprays," ASME *Journal of Fluids Engineering*, Vol. 101, Sept. 1979, pp. 312–318.

3 Alpert, R. L., and Mathews, M. K., "Calculation of Large-Scale Flow Fields Induced by Droplet Sprays,' in *Polyphase Flow and Transport Technology*, Bajura, R. A., Ed., ASME, 1980, pp. 115–127.

4 Heskestad, G., Kung, H. C., and Todtenkopf, N. F., "Air Entrainment Into Water Sprays and Spray Curtains," ASME Paper 76-WA/FE-40, 1976, 12 pages.

5 Rasbash, D. J., and Stark, G. W. V., "Some Aerodynamic Properties of Sprays," *The Chemical Engineer*, No. 164, Dec. 1962, pp. A83–A88.

6 Schlichting, H. *Boundary Layer Theory*, 7th Ed., McGraw Hill, New York, 1979.

7 Singleton, R. E., "The Compressible Gas-Solid Particle Flow Over a Semi-Infinite Flat Plate," *ZAMP Journal of Applied Mathematics and Physics*, Vol. 16, 1965, pp. 421–449.

8 Wallis, G. B., *One Dimensional Two-Phase Flow*, McGraw Hill, New York, 1969.

9 Yeung, W.-S., *Air-Water Spray Analysis*, Topical Report EPRI NP-1879, Electric Power Research Institute, Palo Alto, Calif., 94303, June 1981.

10 Sandoz, S. A., and Sun, K. H., "Modeling Environmental Effects on Nozzle Spray Distribution," ASME paper 76-WA/FE-39, 1976.

11 Chuah, Y. K., Yeung, W.-S., and Fisher, E. N., *Measurement of Air Velocity in a Plane Turbulent Spray*, Report, Thayer School of Engineering, Dartmouth College, Hanover, N.H. 03755, Nov. 1980.

APPENDIX

As an example, we consider the conical spray where the droplet motion is affected by the drag force. Substituting $\sigma = 1$ into equations (17)–(19) and (23), and solving the resulting system, we obtain $p = q = 0$, $m = -1$, and $n = -\frac{1}{2}$. From equation (21), it follows that the shape of the spray envelope is proportional to $x^{-\frac{1}{2}}$. This is highly unrealistic for a gas-liquid spray system. Hence we conclude that no realistic similarity solution of the types (13) and (14) exists for this case.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Strong Stability of Impulsively Initiated Couette Flow for Both Axisymmetric and Non-axisymmetric Disturbances

J. -C. Chen
Graduate Student.

G. P. Neitzel
Assistant Professor.
Mem. ASME

Department of Mechanical and Energy Systems Engineering, Arizona State University, Tempe, Ariz. 85287

A viscous, incompressible fluid is enclosed between two infinitely long coaxial cylinders. The entire system is assumed to be rotating initially as a rigid body with angular velocity Ω. At time t = 0, the outer cylinder is impulsively stopped. The resulting unsteady Couette flow is subject to centrifugal instabilities. Energy theory (strong stability) calculations have been performed for both axisymmetric and nonaxisymmetric disturbances to verify that the most dangerous mode is axisymmetric. The results include global stability bounds, lower bounds on the onset time, and/or upper bounds on the decay times, and are compared to previous marginal stability results for the same basic state.

Introduction

The stability of steady Couette flow between concentric circular cylinders was first investigated, both theoretically and experimentally, by G. I. Taylor [1] in 1923. This problem has continued to receive substantial attention in the period since that time. In recent years, the stability of time-dependent Couette flow has become a subject of intense study. Chen and his coworkers [2–4] have considered the stability of time-dependent Couette flow due to an impulsively started inner cylinder both experimentally and theoretically. Hall [5], Seminara and Hall [6], Riley and Laurence [7], and Tustaniwskyj and Carmi [8] have examined various time-periodic flows between concentric cylinders due to either a modulated azimuthal pressure gradient or modulated cylinder speeds. The former pair of papers [5, 6] employ linear stability theory, while the latter [7, 8] treat their respective problems using energy stability theory. Eagles [9] and Neitzel [10] have examined the problem of circular Couette flow where the inner cylinder angular velocity is a smooth but aperiodic function of time using linear and energy theories, respectively. Tillmann [11] performed experiments for time-dependent Couette flow in which the outer cylinder was impulsively stopped. Neitzel [12] applied energy theory with a marginal stability criterion to impulsively initiated Couette flow and spin-decay, obtaining lower bounds on the onset times for instability.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February, 1982; final revision, May, 1982. Paper No. 82-WA/APM-20.

Copies will be avialable until July, 1983.

The problem examined in this paper is essentially that treated by Tillmann. Consider a viscous, incompressible fluid that is enclosed between two infinitely long coaxial cylinders of radii R_1 and $R_2 > R_1$. The entire system is assumed to be rotating as a rigid body with angular velocity Ω. At time $t = 0$, the angular velocity of the outer cylinder is impulsively brought to zero, while the angular velocity of the inner cylinder is held constant at Ω. Due to the action of viscosity, an unsteady swirl flow is established in the azimuthal direction. Such swirling flows may be subject to centrifugal instability due to an imbalance between the local centrifugal force and pressure gradient acting on the fluid. It is easily shown that the initial state is globally stable. As $t \to \infty$, the flow becomes steady Couette flow, which may be unstable for some values of Reynolds number $R = \Omega(R_2 - R_1)^2/\nu$, where ν is the kinematic viscosity of the fluid, or stable for other values. Prior to the time when this steady flow is essentially reached, disturbances to the time-dependent Couette flow may grow after some "early" time t_o, called the onset time. If the final state is stable, they must then begin to decay after some "long" time t_d, called the decay time. For intermediate times t such that $t_o < t < t_d$, disturbances may grow to possibly large amplitude, but ultimately decay. If the final state is unstable, then instabilities will persist indefinitely after the onset time t_o, although the final Taylor vortex flow is not necessarily identical to the earliest appearing one.

Two different types of stability analyses might be considered: energy theory and linear theory. Energy theory gives a means for determining the stability of either steady or unsteady hydrodynamic systems with respect to disturbances of arbitrary amplitude. Linear theory provides sufficient conditions for instability to infinitesimal disturbances. For unsteady basic states, linear theory in any of its present forms

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(other than direct numerical integration of the governing partial differential equations) is restricted to flows that are either time-periodic or slowly varying. The latter criterion requires that the disturbances grow on a time scale much smaller than that on which the basic state changes. For impulsively driven flows such as the one considered here, Neitzel [12] has pointed out that linear theory is inapplicable due to the fact that near conditions of marginal stability both the basic state and the disturbances to it change on the same diffusive time scale. Therefore, energy theory represents the only analytical tool available for the analysis of such basic states.

Modern usage of the energy method originated with the work of Serrin [13]. Recently, the energy method has been used successfully to analyze time-dependent basic states. Homsy [14] applied energy theory to the problem of impulsively heated and cooled fluid layers. He used a strong stability criterion which required exponential decay of the disturbance kinetic energy to ensure stability and first showed how to determine lower bounds on the onset time and/or upper bounds on the decay time. Neitzel and Davis [15] applied energy theory with a strong stability criterion to spin-decay within a single cylinder. They found bounds on both onset and decay times for such transiently unstable flows. Gumerman and Homsy [16] introduced and utilized the concept of marginal stability, resulting from the work of Davis and von Kerczek [17], to improve the lower bounds obtained from the strong stability analysis. Marginal stability, which is a weaker criterion than strong stability, does not require exponential energy decay, but only that the volume-integrated disturbance kinetic energy is less than its initial value. Patzer and Homsy [18] applied strong and marginal stability criteria to examine the stability of a liquid drop subjected to a step change in boundary temperature or concentration.

In the present work we will use energy theory under a strong stability assumption to obtain the regions of stability in the Reynolds number-time plane for the impulsively initiated Couette flow described earlier. A feature common to previous analyses of unsteady Couette flows is the assumption that the most dangerous disturbances are axisymmetric. This assumption is a natural one to make, based on the results of steady Couette flow computations (Krueger, Gross, and DiPrima [19]) and experimental evidence. One aim of the present work is to verify the correctness of this assumption, at least for a strong stability analysis of one particular unsteady Couette flow. We therefore assume the wave number m in the azimuthal direction varies from $m=0$ to $m=3$, and show that the most dangerous mode is axisymmetric ($m=0$). A second aim is to complement the results of Neitzel [12] by providing upper bounds on the decay times for disturbances, when such bounds exist, and optimizing the global stability bounds for this class of flows. Results are presented for three different radius ratios $\eta \equiv R_1/R_2$, corresponding to the cases treated by Neitzel [12].

Basic State

We seek a parallel flow, unsteady Couette flow solution of the form $(u,v,w,p) = (0, V(r,t), 0, P(r,t))$ where (u,v,w) are the velocities corresponding to the usual cylindrical coordinates (r, θ, z) and p is the pressure. Under this assumption the Navier-Stokes and continuity equations governing the basic state reduce to

$$\rho \frac{V^2}{r} = \frac{\partial P}{\partial r}, \tag{1}$$

and

$$\frac{\partial V}{\partial t} = \nu\left(\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} - \frac{1}{r^2}V\right), \tag{2}$$

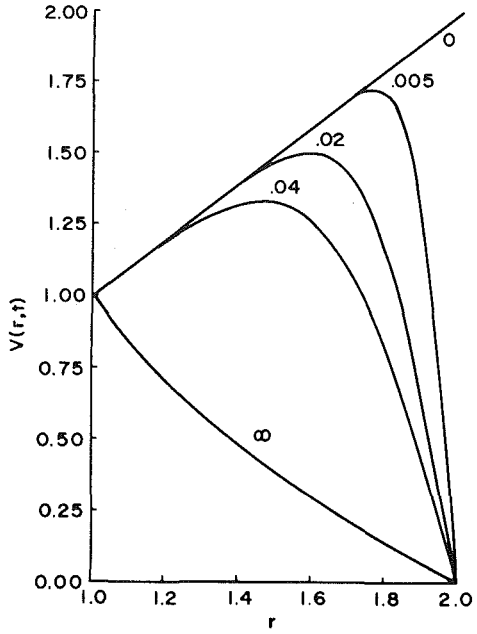


Fig. 1 Azimuthal velocity profiles for impulsive Couette flow ($\eta = 0.5$) at various times

where ρ is the fluid density. The initial condition

$$V(r,0) = r\Omega \tag{3}$$

corresponds to the assumed state of rigid-body rotation with angular velocity Ω while the boundary conditions

$$V(R_1,t) = R_1\Omega \tag{4}$$

$$V(R_2,t) = 0 \tag{5}$$

require no-slip at the solid walls located at $r=R_1, R_2$.

The solution for the velocity field is easily determined [12] using separation of variables and is given by

$$V(r,t) = Ar + B/r + \sum_{n=1}^{\infty} D_n e^{-\lambda_n^2 t}[J_1(\lambda_n r) + C_n Y_1(\lambda_n r)] \tag{6}$$

where J and Y are Bessel functions of the first and second kind, respectively,

$$A = \frac{\eta^2}{\eta^2 - 1}, \tag{7}$$

$$B = \frac{\eta^2}{(1-\eta^2)(1-\eta)^2}, \tag{8}$$

$$C_n = -\frac{J_1\left(\frac{\lambda_n}{1-\eta}\right)}{Y_1\left(\frac{\lambda_n \eta}{1-\eta}\right)} \tag{9}$$

$$Z_k = J_k + C_n Y_k, \tag{10}$$

$$D_n = \frac{2 Z_0\left(\frac{\lambda_n}{1-\eta}\right)}{\lambda_n\left[\eta^2 Z_0^2\left(\frac{\lambda_n \eta}{1-\eta}\right) - Z_0^2\left(\frac{\lambda_n}{1-\eta}\right)\right]}, \tag{11}$$

and λ_n are the roots of

$$J_1\left(\frac{\lambda_n \eta}{1-\eta}\right) Y_1\left(\frac{\lambda_n}{1-\eta}\right) - J_1\left(\frac{\lambda_n}{1-\eta}\right) Y_1\left(\frac{\lambda_n \eta}{1-\eta}\right) = 0. \tag{12}$$

In equations (6)–(12), all quantities have been nondimensionalized using the scales $(R_2 - R_1)$, $(R_2 - R_1)^2/\nu$, and

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

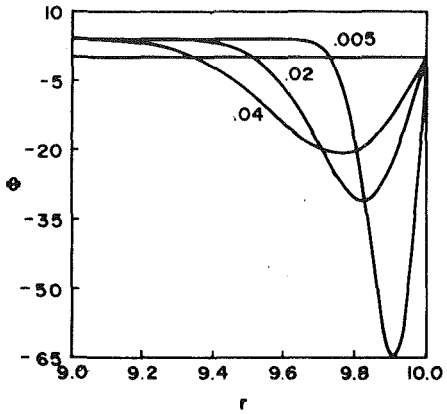


Fig. 2 Rayleigh discriminant versus radius for impulsive Couette flow ($\eta = 0.9$) at various times

$(R_2 - R_1)\Omega$ for length, time, and velocity, respectively, and $\eta = R_1/R_2$ is the radius ratio.

Figure 1 shows instantaneous velocity profiles obtained from (6) for Couette flow ($\eta = 0.5$) at various times. Rayleigh's criterion states that a necessary and sufficient condition for a steady, inviscid, swirl flow $V(r)$ to be linearly stable in the interval $\eta/(1-\eta) \leq r \leq 1/(1-\eta)$ is that the inequality

$$\Phi(r) = \frac{1}{r^3}\frac{d}{dr}|rV(r)|^2 > 0$$

be satisfied for all $r \epsilon [\eta/(1-\eta),\ 1/(1-\eta)]$. The quantity $\Phi(r)$ is called the Rayleigh discriminant. We can calculate the Rayleigh discriminant as a function of radius for *fixed* times for our unsteady flow just to see if one might suspect swirling instabilities to occur. Figure 2 is a plot of instantaneous Φ versus r at various times for $\eta = 0.9$. From this, we can see that according to Rayleigh's criterion, we might expect centrifugal instabilities near the outer wall. Rayleigh's criterion is based on the assumption that the fluid is inviscid and the flow is steady. Since viscosity is a stabilizing quantity for swirling flow, a velocity profile that is predicted to be unstable on the basis of Rayleigh's criterion may indeed be stable in practice due to the effect of viscosity. However, since the Couette flows considered here are both viscous and unsteady, only direct calculation can give stability estimates. The energy method to be applied here gives such estimates for disturbances of arbitrary amplitude.

Stability Analysis

The energy identity is derived in the standard way by taking the inner product of $\mathbf{u} = (u,v,w)$, the disturbance velocity vector, with the usual disturbance equation and then integrating the result over a volume $\forall$ defined by

$$\forall = \{(r,\theta,z) \mid \eta/(1-\eta) \leq r \leq 1/(1-\eta),\ 0 \leq \theta \leq 2\pi,\quad 0 \leq z \leq Z\} \tag{13}$$

where we have assumed axial periodicity with period Z, consistent with an expected Taylor vortex flow. Use of the periodicity in z, Green's theorem, and homogeneous conditions for $\mathbf{u}$ at solid walls leads to the energy identity [13]

$$\frac{dE}{dt} = RI - D \tag{14}$$

where

$$E = \frac{1}{2} <\mathbf{u}\cdot\mathbf{u}> \tag{15}$$

is the kinetic energy of the disturbance,

$$I = <uv(V/r - V_r)> \tag{16}$$

is an energy production term,

$$D = <\nabla \mathbf{u} : \nabla \mathbf{u}> \tag{17}$$

is the viscous dissipation term, and

$$R = \frac{\Omega(R_2 - R_1)^2}{\nu} \tag{18}$$

is the Reynolds number. In equations (15)-(17), $<\ >$ is used to denote the volume integral over $\forall$.

There exists [13, 14] a positive number, $R_E(t)$, such that

$$\frac{1}{R_E(t)} = \max_S \frac{I}{D} \tag{19}$$

where $S = \{\mathbf{u} = (u,v,w) \mid u,v,w \epsilon C^2,\ \mathbf{u} = 0$ at $r = \eta/(1-\eta)$ and $r = 1/(1-\eta),\ \nabla\cdot\mathbf{u} = 0,\ \mathbf{u}$ is periodic in z with period $Z\}$. S is the set of so-called kinematically admissible functions which includes the solutions to the disturbance equations as a subset. For Reynolds numbers $R < R_E(t)$ the flow is strongly stable in the sense that disturbance kinetic energy experiences exponential decay. Hence, $R < R_E(t)$ is a sufficient condition for stability to disturbances of arbitrary amplitude. It is not a necessary condition as well since the members of class S are not required to be solutions to the full dynamical equations.

The Euler-Lagrange equations corresponding to (19) are easily found to be

$$\frac{1}{2}\tilde{R}\Psi v + \frac{1}{r}(r\,u_r)_r + \frac{1}{r^2}u_{\theta\theta} + u_{zz} - \frac{1}{r^2}u - \frac{2}{r^2}v_\theta - \pi_r = 0 \tag{20}$$

$$\frac{1}{2}\tilde{R}\Psi u + \frac{1}{r}(r\,v_r)_r + \frac{1}{r^2}v_{\theta\theta} + v_{zz} - \frac{1}{r^2}v + \frac{2}{r^2}u_\theta - \frac{1}{r}\pi_\theta = 0 \tag{21}$$

$$\frac{1}{r}(rw_r)_r + \frac{1}{r^2}w_{\theta\theta} + w_{zz} - \pi_z = 0 \tag{22}$$

$$\frac{1}{r}(ru)_r + \frac{1}{r}v_\theta + w_z = 0. \tag{23}$$

$\tilde{R}$ and π are Lagrange multipliers with $\tilde{R}$ having the significance that

$$R_E = \min_S \tilde{R}. \tag{24}$$

The corresponding boundary conditions are

$$u = v = w = 0 \tag{25}$$

at $r = 1/(1-\eta)$ and $r = \eta/(1-\eta)$ and the quantity Ψ is defined as

$$\Psi(r,t) = V(r,t)/r - V_r(r,t).$$

Equations (20)–(23) are linear although disturbance amplitude is unrestricted, and time now appears only parametrically through $\Psi(r,t)$.

An interpretation of the types of results obtainable from this analysis may be found elsewhere [14, 15] and will not be repeated here. Suffice it to say that for each case considered, one obtains a curve in the R_E-time plane. For points on an experimental trajectory below this curve, the flow is strongly stable. Therefore, the intersection of an experimental trajectory (a horizontal line at a fixed R) with such a curve will yield either a lower bound on the onset time of an instability, an upper bound on the decay time, or both.

Solution of the Euler-Lagrange Equations

Since the disturbances are periodic in both z and θ, we shall

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

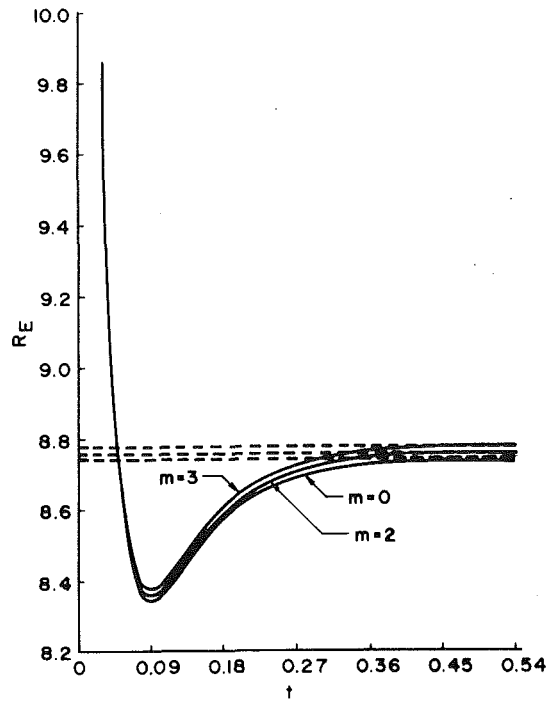


Fig. 3 R_E versus t for $\eta = 0.9$ for various m (large t results)

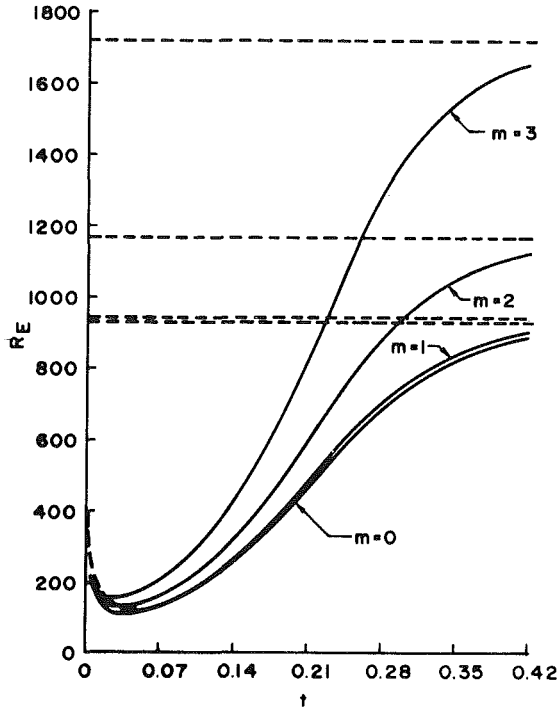


Fig. 5 R_E versus t for $\eta = 0.1$ for various m

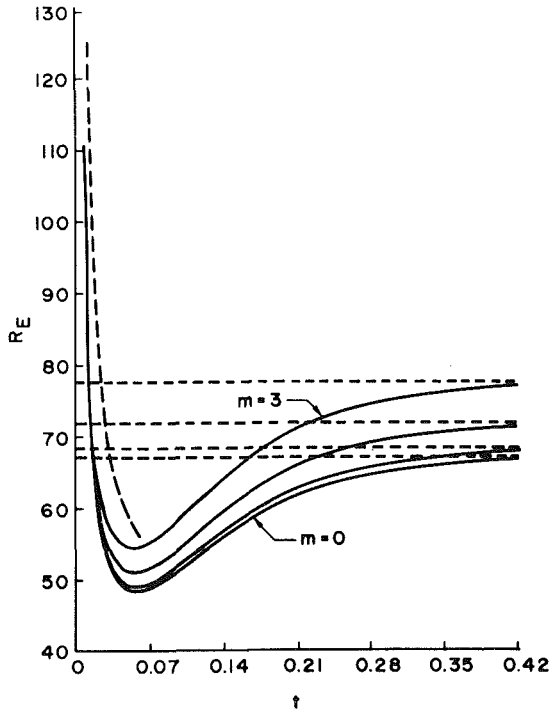


Fig. 4 R_E versus t for $\eta = 0.5$ for various m

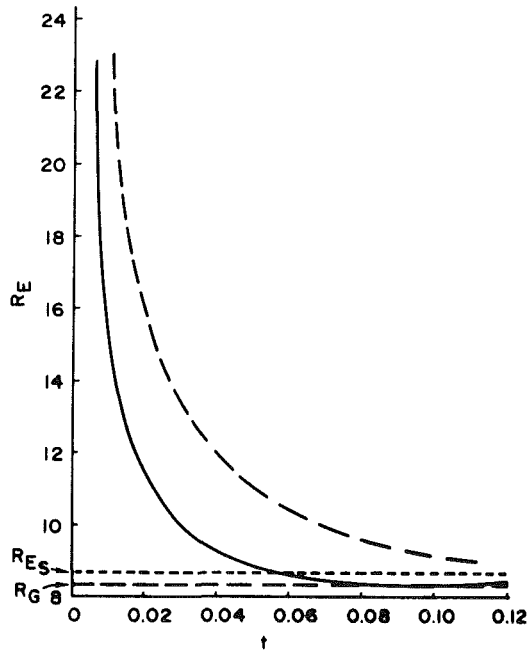


Fig. 6 R_E versus t for $\eta = 0.9$ for $m = 0$ (small t results)

assume that u, v, w, and π may be decomposed into normal modes as

$$u(r,\theta,z;t) = \tilde{u}(r;t)\cos kz\, e^{im\theta} \tag{26}$$

$$v(r,\theta,z;t) = \tilde{v}(r;t)\cos kz\, e^{im\theta} \tag{27}$$

$$w(r,\theta,z;t) = \tilde{w}(r;t)\sin kz\, e^{im\theta} \tag{28}$$

$$\pi(r,\theta,z;t) = \tilde{p}(r;t)\cos kz\, e^{im\theta} \tag{29}$$

where $k = 2\pi/Z$ is the axial wavenumber in accordance with our assumption of axial periodicity. Substituting equations (26)–(29) into the Euler-Lagrange equations (20)–(23) and dropping tildes on $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, and $\tilde{p}$, we get

$$\frac{1}{2}\tilde{R}\Psi v + u'' + \frac{1}{r}u' - \left(\frac{m^2+1}{r^2} + k^2\right)u - \frac{2im}{r^2}v - p' = 0 \tag{30}$$

$$\frac{1}{2}\tilde{R}\Psi u + v'' + \frac{1}{r}v' - \left(\frac{m^2+1}{r^2} + k^2\right)v + \frac{2im}{r^2}u - \frac{im}{r}p = 0 \tag{31}$$

$$w'' + \frac{1}{r}w' - \left(\frac{m^2}{r^2} + k^2\right)w + kp = 0 \tag{32}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

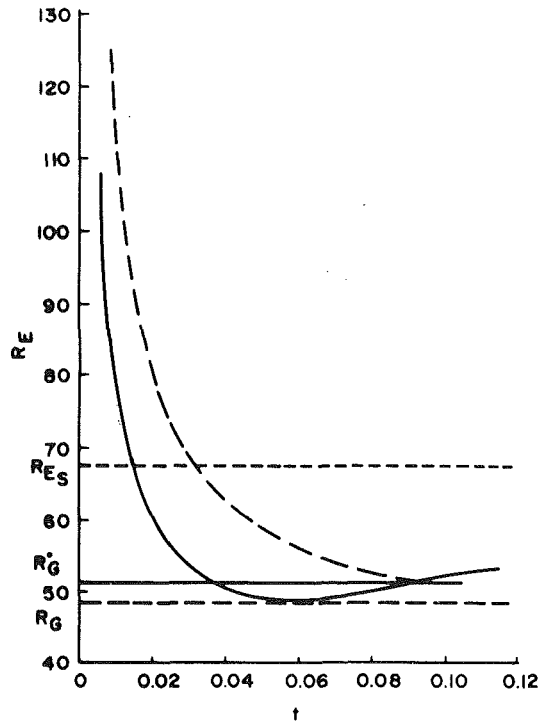


Fig. 7 R_E versus t for $\eta = 0.5$ for $m = 0$ (small t results)

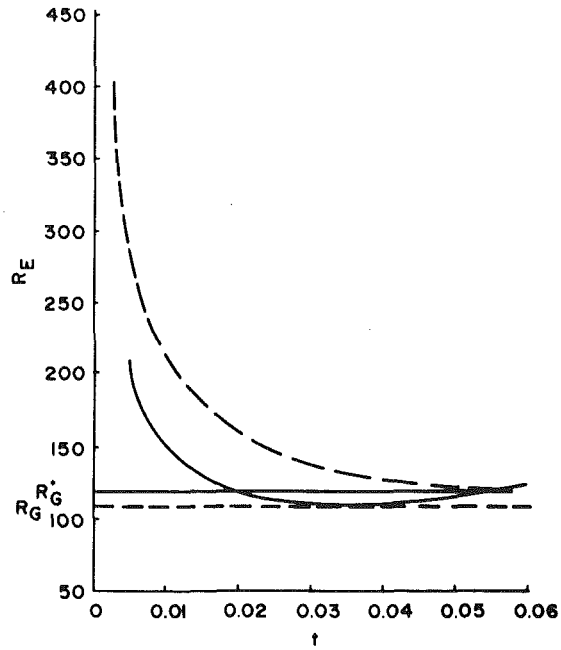


Fig. 8 R_E versus t for $\eta = 0.1$ for $m = 0$ (small t results); $R_{E_S} = 930$

$$u' + \frac{1}{r}u + \frac{im}{r}v + kw = 0 \tag{33}$$

where a prime denotes differentiation with respect to r. The boundary conditions remain

$$u = v = w = 0 \quad \text{at} \quad r = 1/(1-\eta) \quad \text{and} \quad r = \eta/(1-\eta). \tag{34}$$

The procedure for determining R_E at a fixed time t is as follows: for a fixed t, eigenvalues $\tilde{R}$ of system (30)–(34) are determined for various k and m, where m must be in integer to guarantee single-valuedness of the physical quantities. Then, the minimum problem of equation (24) becomes

$$R_E = \min_{k,m} \tilde{R}. \tag{35}$$

Eigenvalues $\tilde{R}$ were determined for fixed m and k using a standard shooting procedure without orthonormalization. Three-point Lagrangian interpolation was used to determine the zeros of the resulting determinant as well as the critical value of k at which a minimum occurred for each m. Since the equations are complex, the determinant is likewise complex. Once a zero of the real part of the determinant was located, the imaginary part was checked to insure that it also changed sign. The azimuthal wave number m was varied from 0 to 3 to see which modes were the least stable. Details of the calculation procedure and tabulated results are given by Chen [20].

Results, Conclusions, and Discussion

The calculations described in the preceding section were performed in double precision using an Amdahl 470 V7B computer for three different cases of Couette flow ($\eta = 0.9$, 0.5, 0.1). The results are shown in Figs. 3–8 in terms of Reynolds number versus time. Shown also on each graph is the value of R_{E_S} corresponding to the energy limit for the steady-state flow which is approached as time becomes large.

The effect of the azimuthal wavenumber m is of interest. Figures 3, 4, and 5 are plots of R_E versus time as a function of the wavenumber m. The small time, high Reynolds number values have been omitted from Fig. 3 to allow the vertical scale to be expanded so that the curves corresponding to various modes may be seen. Also omitted from Fig. 3 is the curve corresponding to $m = 1$, which is very close to, but above, the $m = 0$ curve. The small time results for this case appear on Fig. 6. The curves show the expected result that the most dangerous disturbances are axisymmetric ($m = 0$). At any fixed time, the critical value of R_E for an axisymmetric disturbance is always less than the critical values of R_E for all nonaxisymmetric disturbances. Therefore, axisymmetric disturbances have the smallest lower bounds on onset times and the largest upper bounds on decay times.

For the cases of $\eta = 0.9$, 0.5, and 0.1 (Figs. 3, 4, and 5, respectively) a global stability result, R_G, can be obtained by estimating the value of R_E at the minimum point on the curve for an axisymmetric disturbance ($m = 0$). This would appear to be around $R_G = 8.3$ for $\eta = 0.9$, $R_G = 48.3$ for $\eta = 0.5$, and $R_G = 109.2$ for $\eta = 0.1$. The curves of $m = 0$ show lower bounds on the onset time t_o and the upper bounds on the decay time t_d obtainable for each $R_G < R < R_{E_S}$, or only lower bounds on the onset time for each $R > R_{E_S}$. As $t \to 0$, the computations become more difficult because more and more terms are required to accurately represent the basic state variable, Ψ, and gradients in the shooting determinant with respect to $\tilde{R}$ become very large. However, the computed curve should be monotonic in the range of small t and large R.

Neitzel [12] has performed marginal stability computations for the Couette flow considered here as well as for the related problem of spin-decay. These computations were for axisymmetric disturbances. The results are plotted on Figs. 4–8 as dashed curves. A comparison between the strong and marginal stability results shows the expected shift to the right of the marginal stability curve. Therefore, the onset time bounds predicted by the marginal stability analysis are less conservative than those of the strong stability analysis. Marginal stability, however, does not yield meaningful results regarding decay times. This information comes only from a strong stability analysis. Enhanced global stability results are obtainable [12] by examining the marginal and strong stability results together. Figures 6–8 show the results for the three cases treated emphasizing the small time results for $m = 0$. In Fig. 6, the strong stability curve dips below the value R_{E_S} indicating the possibility of transient instability. The marginal stability curve, however, approaches R_{E_S} monotonically from

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

above, indicating that transient instability will not be seen in this "narrow" gap case. Therefore, the appropriate global stability limit here is R_{E_S}. In Figs. 7 and 8, both analyses indicate possible transient instability below R_{E_S}. In these cases, an improved global stability result R_G^* may be determined from the intersection of the marginal and strong stability curves. For an experimental trajectory below this value, the flow will be marginally stable for all times prior to the time where the curves intersect, and then strongly stable for all subsequent times. Therefore, the quantity $E(t)$ (equation (15)) will be bounded for all time by its initial value. For the case of $\eta = 0.5$ the optimal global stability bound $R_G^* = 51$ is 6 percent higher than the strong stability value of $R_G = 48$. For the case of $\eta = 0.1$, $R_G^* = 120$ is 10 percent higher than the strong stability value of $R_G = 109$. These are in rough agreement with those estimated [12] from the marginal stability results alone.

In conclusion, it has been known for some time [19] that for steady Couette flow between rotating cylinders (when both rotate in the same direction), axisymmetric modes are the most unstable. The present analysis has shown that for the particular unsteady flow under consideration, these modes are least stable under an assumption of strong stability. This result does not necessarily imply that an axisymmetric disturbance assumption is correct for the previous marginal stability work of reference [12], or for other types of time dependence, but is does give one some degree of faith that this may indeed be the case. In addition to investigating this assumption, we have provided global stability bounds and upper bounds on decay times for the cases examined that complement the onset time bounds obtained from marginal stability.

Acknowledgment

This work was supported by the Fluid Mechanics Program of the National Science Foundation under Grant No. CME-8006696 and, in part, by an Arizona State University Faculty Grant-In-Aid.

References

1 Taylor, G. I., "Stability of a Viscous Liquid Contained Between Two Rotating Cylinders," *Phil. Trans. Roy. Soc. A*, Vol. 223, 1923, pp. 289–343.

2 Kirchner, R. P., and Chen, C. F., "Stability of Time-Dependent Rotational Couette Flow. Part 1. Experimental Investigation," *J. Fluid Mech.*, Vol. 40, 1970, pp. 39–47.

3 Chen, C. F., and Kirchner, R. P., "Stability of Time-Dependent Rotational Couette Flow. Part 2. Stability Analysis," *J. Fluid Mech.*, Vol. 48, 1971, pp. 365–384.

4 Liu, D. C. S., and Chen, C. F., "Numerical Experiments on Time-Dependent Rotational Couette Flow," *J. Fluid Mech.*, Vol. 59, 1973, pp. 77–95.

5 Hall, P., "The Stability of Unsteady Cylinder Flows," *J. Fluid Mech.*, Vol. 67, 1975, pp. 29–63.

6 Seminara, G., and Hall, P., "Linear Stability of Slowly Varying Unsteady Flows in a Curved Channel," *Proc. Roy. Soc. London A*, Vol. 346, 1975, pp. 279–303.

7 Riley, P. J., and Laurence, R. L., "Energy Stability of Modulated Circular Couette Flow," *J. Fluid Mech.*, Vol. 79, 1977, pp. 535–552.

8 Tustaniwskyj, J. I., and Carmi, S., "Nonlinear Stability of Modulated Finite Gap Taylor Flow," *Phys. Fluids*, Vol. 23, 1980, pp. 1732–1739.

9 Eagles, P. M., "On the Stability of Slowly-Varying Flow Between Concentric Cylinders," *Proc. Roy. Soc. London A*, Vol. 355, 1977, pp. 209–224.

10 Neitzel, G. P., "Stability of Circular Couette Flow With Variable Inner Cylinder Speed," *J. Fluid Mech.*, Vol. 123, 1982, pp. 43–57.

11 Tillmann, W., "Development of Turbulence During the Build-Up of a Boundary Layer at a Concave Wall," *Phys. Fluids Supp.*, Vol. 10, 1967, pp. S108–S111.

12 Neitzel, G. P., "Marginal Stability of Impulsively Initiated Coutte Flow and Spin-Decay," *Phys. Fluids*, Vol. 25, 1982, pp. 226–232.

13 Serrin, J., "On the Stability of Viscous Fluid Motions," *Arch. Rat. Mech. Anal.*, Vol. 3, 1959, pp. 1–13.

14 Homsy, G. M., "Global Stability of Time-Dependent Flows: Impulsively Heated or Cooled Fluid Layers," *J. Fluid Mech.*, Vol. 60, 1973, pp. 129–139.

15 Neitzel, G. P., and Davis, S. H., "Energy Stability Theory of Decelerating Swirl-Flows," *Phys. Fluids*, Vol. 23, 1980, pp. 432–437.

16 Gumerman, R. J., and Homsy, G. M., "The Stability of Uniformly Accelerated Flows With Application to Convection Driven by Surface Tension," *J. Fluid Mech.*, Vol. 68, 1975, pp. 191–207.

17 Davis, S. H., and von Kerczek, C., "A Reformulation of Energy Stability Theory," *Arch. Rat. Mech. Anal.*, Vol. 52, 1973, pp. 112–117.

18 Patzer, J. F., II, and Homsy, G. M., "Global Stability of Transient Drop Extraction to Marangoni Instabilities," *Phys. Fluids*, 1981, Vol. 24, pp. 567–575.

19 Krueger, E. R., Gross, A., and DiPrima, R. C., "On the Relative Importance of Taylor-Vortex and Non-Axisymmetric Modes in Flow Between Rotating Cylinders," *J. Fluid Mech.*, Vol. 24, 1966, pp. 521–537.

20 Chen, J. -C., "Strong Stability of Impulsively Initiated Couette Flow for Both Axisymmetric and Non-Axisymmetric Disturbances," MS Thesis, Department of Mechanical and Energy Systems Engineering, Arizona State University, 1982.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Model of the Turbulent Burst Phenomenon: Fully Developed Transitional Turbulent Flow Between Parallel Plates

L. C. Thomas

H. M. Kadry

Department of Mechanical Engineering,
University of Petroleum
and Minerals,
Ministry of Higher Education,
Dhahran, Saudi Arabia

An analysis is presented for fully developed transitional turbulent flow between parallel plates that features the use of a turbulent burst model for the important wall region. Model closure is accomplished by the specification of the mean burst frequency and, for moderate to high Reynolds numbers, by matching with a classical eddy viscosity model for the turbulent core. Predictions obtained for friction factor and mean velocity distribution are compared with experimental data for fully developed transitional turbulent flow in a channel with large aspect ratio. Predictions are also developed for the eddy viscosity within the wall region for transitional turbulent conditions.

Introduction

Whereas fully turbulent internal flows have been the subject of numerous studies over the past 20 years, relatively little attention has been given to transitional turbulent flows. Unlike fully turbulent flow fields, transitional turbulent flow is characterized by dimensionless velocity distributions u^+ ($\equiv \bar{u}/u_\tau$) that are strongly dependent on Reynolds number Re.

Pai [1] and Wilson and Azad [2] were among the first to analyze this problem. Both of these analyses are based on the idea that the overall transport process for transitional flow consists of laminar and turbulent components. Pai assumed that some fraction of the mean velocity distribution must be parabolic, and Wilson and Azad assumed that some fraction of the eddy diffusivity for momentum ϵ_m must be equal to the molecular kinematic viscosity ν, with the remaining portion equal to the values of ϵ_m associated with fully turbulent flow.

Gill and Scher [3] developed an analysis for transitional turbulent flow between parallel plates involving an intermittency factor γ_t of the form $1-\exp(-\phi/w/2)$, where $\phi = (\mathrm{Re}-a)/b$, and a and b are empirical constants. In this approach, the mixing length l for fully turbulent conditions is multiplied by γ_t. However, no theoretical or experimental justification was offered in support of this intermittency factor.

One and two-equation transport equation turbulence models have also been utilized in analyzing transitional turbulent flows [4–6], with some degree of success. For example, Jones and Launder [4, 5] have developed a $K-\epsilon$ model for transitional turbulent flow between parallel plates. However, to account for the effects of low turbulent Reynolds number, rather arbitrary ad hoc assumptions (including damping factor corrections) were employed in both the mean turbulent kinetic energy K and dissipation ϵ equations. As stated by the developers of turbulence models of this type, these assumptions for modeling wall turbulence still require further adjustments and refinements.

The main difficulty encountered in utilizing classical eddy viscosity or higher order approaches in analyzing transitional turbulent flow involves the lack of information pertaining to ϵ_m and other turbulence characteristics in the wall region. Because of deficiencies in the traditional eddy diffusivity approach to transitional turbulent convection heat transfer, Lawn [7] has called for the development of alternative modeling concepts. Others such as Spalding [8] and Kays [9] have recognized the lack of a fundamental model of wall turbulence as one of the limitations in the current approaches to analyzing turbulent convection transport processes in general.

In this regard, an alternative approach to analyzing transitional turbulent flow has been introduced which involves modeling the turbulent burst phenomenon [10]. In this surface-renewal approach, the actual unsteady process that occurs within coherent large-scale structures between the inrush and ejection phases of a burst event is modeled. Similar to earlier surface-renewal type models for fully turbulent flow by Einstein and Li [11], Hanratty [12], Thomas [13], and others, a preliminary approach was used in affecting model closure that involved a rather crude representation of the overlying turbulent core. Nevertheless, predictions developed for u^+ within the wall region and for the mean frequency $\bar{s}$ of the burst process were found to be in reasonable agreement

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 1981; final revision, December 1981. Paper No. 82-WA/APM-10.

Copies will be available until July, 1983.

Copyright © 1982 by ASME

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with experimental data. These positive findings were felt to demonstrate the value of the basic surface-renewal modeling concept. However, the analysis of reference [10] involved the use of empirical information for the friction factor f, and the predictions for u^+ were found to fail in the zone immediately outside the wall region, especially for increasing values of Reynolds number. The objective of the present paper is to improve the surface-renewal approach to analyzing transitional turbulent flow by specifying the mean burst frequency $\bar{s}$ rather than f, and by matching the surface-renewal model for wall turbulence with an eddy diffusivity model for the turbulent core.

Analysis

Following the pattern established in reference [14], the instantaneous transfer of axial momentum within an individual large-scale structure residing within the wall region between inrush and ejection is represented by the unsteady three-dimensional Navier-Stokes equation,

$$\frac{\partial u}{\partial \theta} + u_j \frac{\partial u}{\partial x_j} = \nu \frac{\partial^2 u}{\partial x_j \partial x_j} - \frac{1}{\rho}\frac{\partial P}{\partial x} \tag{1}$$

where the contact time θ is equal to zero at the instant of inrush. The initial and boundary conditions are given by

$$u = U_i \quad \text{at} \quad \theta = 0 \tag{2}$$

$$u = 0 \quad \text{at} \quad y = 0 \tag{3}$$

$$u = u_M \quad \text{at} \quad y = y_M \tag{4}$$

plus additional streamwise and spanwise boundary conditions; U_i is the mean axial velocity of fluid brought into the wall region by the inrush process, and u_M is the instantaneous velocity at the edge of the wall region and the turbulent core.

The contribution of the instantaneous transport within the many individual elements of fluid residing within the wall region at any instant of process time to the spatial average transport is accounted for by the use of the statistical age distribution concept. Based on this approach, a mean transport characteristic $\bar{\psi}$ (i.e., $\bar{u}$ or $\bar{P}$) is expressed in terms of the instantaneous characteristic ψ by

$$\bar{\psi} = \int_0^\infty \psi \phi(\theta, \bar{s})\, d\theta \tag{5}$$

where $\bar{s}$ is the mean frequency of the turbulent burst process. The product $\phi(\theta, s)d\theta$ represents the fraction of the wall region fluid with age between θ and $\theta + d\theta$. Although experimental measurements for $\phi(\theta, \bar{s})$ are not presently available for turbulent boundary layer flows, Davies and Lozano [15] have recently reported experimental evidence that indicates that the rates of replacement of eddies at the surface of a stirred vessel by fresh eddies is consistent with the Danckwerts [16] exponential distribution.

$$\phi(\theta, \bar{s}) = \bar{s} \exp(-\theta \bar{s}) \tag{6}$$

Because of this and because previous studies have indicated that predictions for $\bar{u}$ are fairly insensitive to the form of $\phi(\theta, \bar{s})$, the Danckwerts distribution is used in this analysis.

Utilizing equations (5) and (6), the instantaneous equations are transformed into the mean domain for an assumed uniform distribution in U_i.

$$\bar{s}(\bar{u} - U_i) + \overline{u_j \frac{\partial u}{\partial x_j}} = \nu \frac{\partial^2 \bar{u}}{\partial x_j \partial x_j} - \frac{1}{\rho}\frac{\partial \bar{P}}{\partial x} \tag{7}$$

$$\bar{u} = 0 \quad \text{at} \quad y = 0 \tag{8}$$

$$\bar{u} = \bar{u}_M \quad \text{at} \quad y = y_M \tag{9}$$

As pointed out in reference [14], the term $\bar{s}(\bar{u} - U_i)$ accounts for the important eddy transport associated with the inrush and ejection mechanism. The term $\overline{u_j \partial u/\partial x_j}$ (together with similar terms that appear in the y and z-component momentum equations) is associated with the three-dimensional convective interaction between individual coherent structures and the surrounding fluid. As a first approximation, this convective interaction term is approximated by $\bar{u}_j\, \partial \bar{u}/\partial x_j$, such that equation (7) reduces to the following form for a two-dimensional mean flow field:

$$\bar{s}(\bar{u} - U_i) + \bar{u}\frac{\partial \bar{u}}{\partial x} + \bar{v}\frac{\partial \bar{u}}{\partial y} = \nu\left(\frac{\partial^2 \bar{u}}{\partial x^2} + \frac{\partial^2 \bar{u}}{\partial y^2}\right) - \frac{1}{\rho}\frac{\partial \bar{P}}{\partial x} \tag{10}$$

Using the usual boundary layer approximations $\partial^2 \bar{u}/\partial x^2 \simeq 0$ and $\partial \bar{P}/\partial x \simeq d\bar{P}/dx$, and setting $\partial \bar{u}/\partial x = 0$ and $\bar{v} = 0$ for fully developed conditions, equation (10) takes the form

$$\bar{s}(\bar{u} - U_i) = \nu \frac{d^2 \bar{u}}{dy^2} - \frac{1}{\rho}\frac{\partial \bar{P}}{dx} \tag{11}$$

To close the analysis, these equations must be solved for $\bar{u}$ with the parameters $\bar{s}$, U_i, $\bar{u}_M$, and y_M specified.

As the Reynolds number Re falls toward the transition value (about 2000), the mean frequency of the burst process $\bar{s}$ approaches zero, and equations (8), (9), and (11) reduce to the appropriate form for laminar flow with $\bar{u}_M = 0$ and $y_M = w$. On the other hand, for high Re conditions the extent of wall region is small such that the effect of axial pressure gradient on momentum transfer within the wall layer becomes very small. For such fully turbulent conditions, equation (11) reduces to

$$\bar{s}(\bar{u} - U_i) = \nu \frac{d^2 \bar{u}}{dy^2} \tag{12}$$

The solution to equation (8), (9), and (12) has been matched with the familar logarithmic law for u^+ in the turbulent core to specify the modeling parameters for fully turbulent flow, with the result $u_\tau/\sqrt{\nu \bar{s}} = 15.95$, $U_i/u_\tau = 15.94$, $u_M^+ = 15.76$, and $y_M^+ = 58.17$. An approximate solution has also been obtained for fully turbulent flow which gives nearly the same results by replacing equation (9) by the simplifying boundary condition $\bar{u} = U_i$ (or $\partial \bar{u}/\partial y = 0$) as $y \to \infty$ (17).

Returning to the formulation for transitional turbulent conditions, predictions will be developed for mean velocity distribution $\bar{u}$ and friction factor f in terms of the single modeling parameter $\bar{s}$ for both low Re and moderate to high Re transitional turbulent conditions. The parameters U_i, $\bar{u}_M$, and y_M will be eliminated by matching the wall model with a classical eddy diffusivity model for the turbulent core and by making use of the Newton law of viscous shear. The analysis will be closed by establishing a practical relationship for $\bar{s}$ for transitional turbulent conditions.

Based on the study by Patel and Head [18], the turbulent bursts fill the entire cross section for low Re transitional turbulent conditions, such that the viscous wall region essentially extends across the entire flow field. For example, the dimensionless distance w^+ ($\equiv w u_\tau/\nu$) between two plates is about 100 for Re equal to 3000. It follows that for low Re transitional turbulent conditions the flow field is represented by equations (8), (9), and (11) with $\bar{u}_M = 0$ and $y_M = w$. On the other hand, for values of Re larger than about 4000, the flow field consists of a turbulent core as well as the wall region. For this case, the modeling equations are restricted to the region close to the wall for which $y_M < w/2$. Attention if first focused on low Re transitional turbulent flow and then on transitional turbulent flow on the moderate to large Re range.

Low Re Transitional Turbulent Flow

The solution to equations (8), (9), and (11) with $\bar{u}_M = 0$ and $y_M = w$ is given by

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\bar{u} = \left(U_i - \frac{1}{\bar{s}\rho}\frac{d\bar{P}}{dx}\right)\left[1 - \frac{\cosh(\sqrt{s^*}\,Y(w))}{\cosh(\sqrt{s^*}/2)}\right] \tag{13}$$

where $s^* = w^2\bar{s}/\nu$ and $Y = w/2 - y$. The pressure gradient is related to the mean wall shear stress $\bar{\tau}_0$ for fully developed flow by $d\bar{P}/dx = -2\rho u_\tau^2/w$. Equation (13) expresses the mean velocity distribution $\bar{u}$ in terms of u_τ and the modeling parameters $\bar{s}$ and U_i.

By introducing the Newton law of viscous shear,

$$\bar{\tau}_0 = \mu\frac{\partial\bar{u}}{\partial y}\bigg|_0 = -\mu\frac{\partial\bar{u}}{\partial Y}\bigg|_{w/2} \tag{14}$$

the following independent equation relating u_τ, s^*, and U_i is obtained from equation (13):

$$u_\tau^2 = \left(U_i + \frac{2u_\tau^2 w}{s^*\nu}\right)\nu\frac{\sqrt{s^*}}{w}\tanh\frac{\sqrt{s^*}}{2} \tag{15}$$

The defining equation for bulk stream velocity U_b provides an additional independent relationship.

$$U_b = \frac{1}{w}\int_0^w \bar{u}\,dy$$

$$= \left(U_i + \frac{2u_\tau^2 w}{s^*\nu}\right)\left(1 - \frac{2}{\sqrt{s^*}}\tanh\frac{\sqrt{s^*}}{2}\right) \tag{16}$$

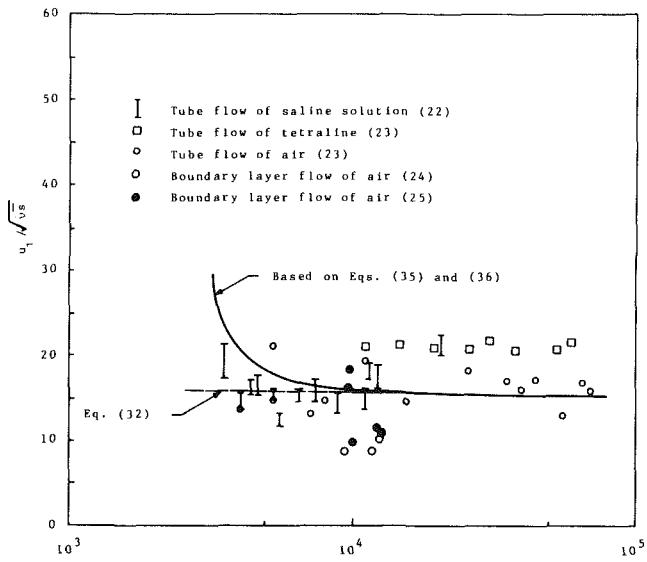


Fig. 1 Dimensionless mean periodicity of wall turbulence fluctuations

By coupling equations (15) and (16), U_i is found to be equal to U_b.

By combining equations (13) and (16), $\bar{u}$ is given by

$$\frac{\bar{u}}{U_b} = \left(1 - \frac{2}{\sqrt{s^*}}\tanh\frac{\sqrt{s^*}}{2}\right)^{-1}\left[1 - \frac{\cosh(\sqrt{s^*}\,Y/w)}{\cosh(\sqrt{s^*}/2)}\right] \tag{17}$$

The dimensionless velocity distribution u^+ is obtained from the relationship

$$u^+ = \frac{\bar{u}}{u_\tau} = \frac{\bar{u}/U_b}{\sqrt{f/2}} \tag{18}$$

where f is the Fanning friction factor. Alternatively, U_i^+ can be eliminated by combining equations (13) and (15), with the result

$$u^+ = \frac{w^+}{\sqrt{s^*}}\coth\frac{\sqrt{s^*}}{2}\left[1 - \frac{\cosh(\sqrt{s^*}\,Y^+/w^+)}{\cosh(\sqrt{s^*}/2)}\right] \tag{19}$$

where $w^+ = \mathrm{Re}\sqrt{f/8}$.

To obtain an expression for f, U_i is set equal to U_b in equation (15).

$$f = \frac{2s^*/\mathrm{Re}}{\frac{\sqrt{s^*}}{2}\coth\frac{\sqrt{s^*}}{2} - 1} \tag{20}$$

The Reynolds number $\mathrm{Re}(\equiv 2wU_b/\nu)$ is based on the hydraulic diameter $2w$ and s^* is to be specified.

Moderate to High Re Transitional Turbulent Flow

In analyzing turbulent momentum transfer for moderate to high Re transitional turbulent flow, mathematical solutions are developed for both the wall region and the turbulent core.

Wall Region. Following the pattern established in the analysis of wall turbulence for fully turbulent flow, a convenient approximate solution to equations (8), (9), and (11) can be obtained for the wall region by replacing equation (9) by

$$\frac{\partial\bar{u}}{\partial y} = 0 \quad \text{as} \quad y \to \infty \tag{21}$$

Because the center line lies outside the wall region for moderate to high Re transitional turbulent flow, the boundary condition $\partial\bar{u}/\partial y = 0$ at $y = w/2$ is equivalent to equation (21). It follows that the dimensionless mean velocity distribution u^+ within the wall region for moderate to high Re transitional turbulent flow can be approximated by equation (19), with the

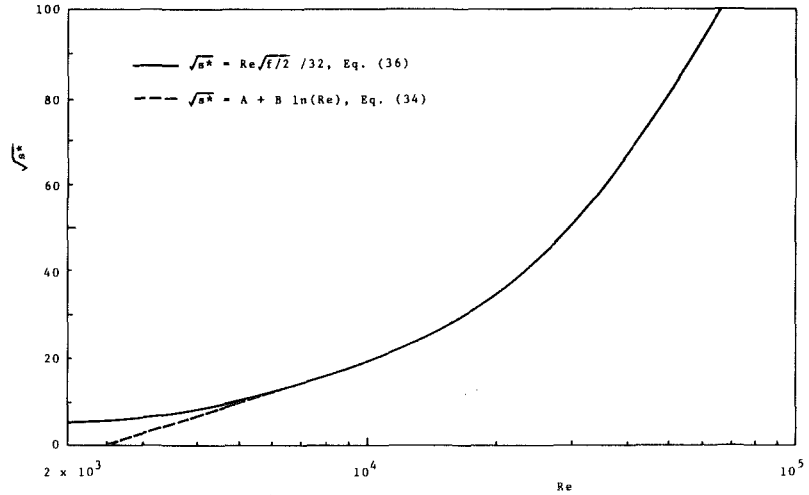


Fig. 2 Dimensionless mean burst frequency s*

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

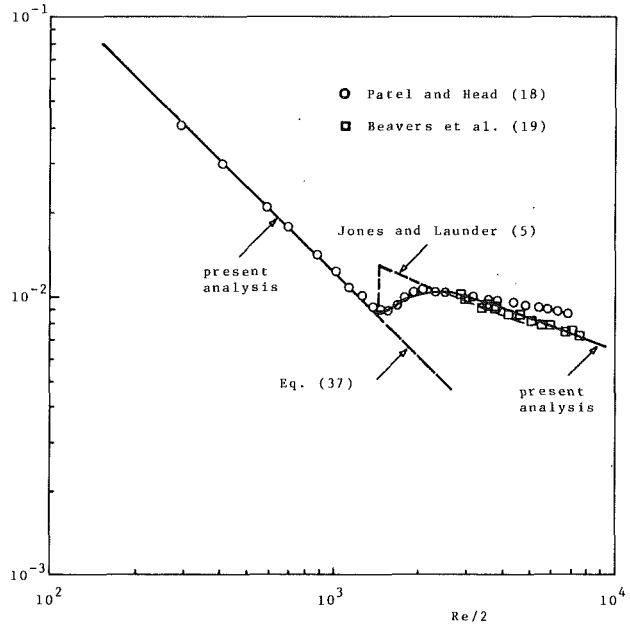


Fig. 3 Fanning friction factor for fully developed transitional turbulent channel flow

relationship between U_i^+ and s^* given by equation (15). To evaluate the friction factor f and to extend the analysis beyond the wall region, equation (19) is coupled with a classical law for u^+ within the turbulent core.

Turbulent Core. The standard eddy viscosity approach provides a means of establishing laws for u^+ in the region away from the wall. Based on this approach, u^+ is expressed in terms of ϵ_m by

$$u^+ = \int \frac{1 - y^+/w^+/2}{1 + \epsilon_m/\nu}\, dy^+ + C_1 \tag{22}$$

One of the simplest and useful correlations for ϵ_m within the turbulent core for fully turbulent flow is given by (9)

$$\frac{\epsilon_m}{\nu} = \rho^{+2}\frac{\partial u^+}{\partial y^+} = (\kappa y^+)^2 \frac{\partial u^+}{\partial y^+} \quad \text{intermediate region}$$

$$y_M^+ \ll y^+ \ll y_1^+ \tag{23a}$$

$$\frac{\epsilon_m}{\nu} = \alpha_1 \delta^+ \quad \text{outer region}$$

$$y_1^+ \ll y^+ \leqslant \delta \tag{23b}$$

where $\kappa \simeq 0.4$, $\alpha_1 \simeq 0.08$, $y_1/\delta \simeq 0.2$, and $\delta = w/2$ for flow between parallel plates. This correlation will also be assumed to apply for high Re transitional turbulent conditions. As Re falls below about 30,000, the fraction of the flow field occupied by the wall region increases. This increase is primarily at the expense of the intermediate region ($\rho^+ = \kappa y^+$), which shrinks as the flow moves toward the transitional turbulent state. For Re between about 10,000 and 4000, the flow field essentially consists of the wall region and the outer region, with the wall region becoming more and more dominant as Re decreases. For this case, ϵ_m within the turbulent core is assumed to be approximated by an expression of the same form as equation (23*b*); i.e.,

$$\frac{\epsilon_m}{\nu} = \alpha_1 \delta^+ \quad y_M^+ \ll y^+ \ll \delta^+ \tag{24}$$

Based on equations (22)-(24), laws are obtained for u^+ within the turbulent core of the following forms:

Moderate Re Transitional Turbulent Flow

$$u^+ = \frac{y^+ - y^{+2}/w^+}{1 + \alpha_1 w^+/2} + C_1 \quad y_M^+ \ll y^+ \ll w^+/2 \tag{25}$$

High Re Transitional Turbulent Flow

$$u^+ = \frac{2}{\kappa}\sqrt{1 - \frac{y^+}{w^+/2}} + \frac{1}{\kappa} ln\left(\frac{4}{P^+}\frac{\sqrt{1 + P^+ y^+} - 1}{\sqrt{1 + P^+ y^+} + 1}\right) + C_1$$

$$y_M^+ < y^+ < y_1^+ \tag{26a}$$

$$u^+ = \frac{y^+ - y^{+2}/w^+}{1 + \alpha_1 w^+/2} + C_2 \quad y_1^+ < y^+ < w^+/2 \tag{26b}$$

where $y_1^+ = 0.1 w^+$ and $P^+ = -1/w^+/2$. C_2 can be eliminated from equation (26*b*) by equating equations (26*a*, *b*) at y_1^+.

To express y_M^+ and C_1 in terms of w^+ (or f) and s^*, equation (19) is matched with equation (25) for moderate Re and with equation (26*a*) for high Re by requiring continuity in u^+ and $\partial u^+/\partial y^+$. It follows that for moderate Re conditions

$$\frac{\sinh[(\sqrt{s^*}/2)(1 - y_M^+/w^+/2)]}{\sinh(\sqrt{s^*}/2)} = \frac{1 - y_M^+/w^+/2}{1 + \alpha_1 w^+/2} \tag{27}$$

$$C_1 = \frac{w^+}{\sqrt{s^*}} \coth\frac{\sqrt{s^*}}{2}\left\{1 - \frac{\cosh[(\sqrt{s^*}/2)(1 - y_M^+/w^+/2)]}{\cosh(\sqrt{s^*}/2)}\right\}$$

$$- \frac{y_M^+ - y_M^{+2}/w^+}{1 + \alpha_1 w^+/2} \tag{28}$$

and for high Re conditions

$$\frac{\sinh[(\sqrt{s^*}/2)(1 - y_M^+/w^+/2)]}{\sinh(\sqrt{s^*}/2)} = \frac{\sqrt{1 - y^+/w^+/2}}{1 + \kappa y^+} \tag{29}$$

$$C_1 = \frac{w^+}{\sqrt{s^*}} \cosh\frac{\sqrt{s^*}}{2}\left\{1 - \frac{\cosh[(\sqrt{s^*}/2)(1 - y_M/w^+/2)]}{\cosh(\sqrt{s^*}/2)}\right\}$$

$$- \frac{2}{\kappa}\sqrt{1 - \frac{y_M^+}{w^+/2}} - \frac{1}{\kappa} ln\left(\frac{4}{P^+}\frac{\sqrt{1 + P^+ y_M^+} - 1}{\sqrt{1 + P^+ y_M^+} + 1}\right) \tag{30}$$

A third independent relationship between C_1, y_M^+, w^+ (or f), and s^* is obtained by using the defining equation for bulk stream velocity U_b; i.e.,

$$U_b = \frac{1}{w}\int_0^w \bar{u}\, dy$$

or

$$\sqrt{\frac{2}{f}} = \frac{\mathrm{Re}/2}{w^+} = \int_0^1 u^+ \mathrm{d}\left(\frac{y^+}{w^+/2}\right) \tag{31}$$

where u^+ is given by equations (19) and (25) for moderate Re conditions and by equations (19) and (26) for high Re conditions. With C_1 given by equations (28) and (30), and with w^+ expressed in terms of $y_M^+/w^+/2$ and s^* by equation (27) and (29), equation (31) gives rise to rather involved implicit expressions for $y_M^+/w^+/2$ in terms of s^* and Re for both moderate and high Re transitional turbulent flows [19]. Thus, with s^* specified in terms of Re, predictions can be obtained for $y_M^+/w^+/2$, w^+, f, C_1 and, for high Re conditions, for C_2 and y_1^+.

Mean Burst Frequency for Transitional Turbulent Flow

Attention is now turned to the specification of the dimensionless mean burst frequency s^*.

Although no experimental data have been reported for mean burst frequency $\bar{s}$ for low Re transitional turbulent channel flow, data are available for moderate to high Re transitional turbulent and fully turbulent fully developed flow in tubes. Representative data for $\bar{s}$ which were obtained for

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

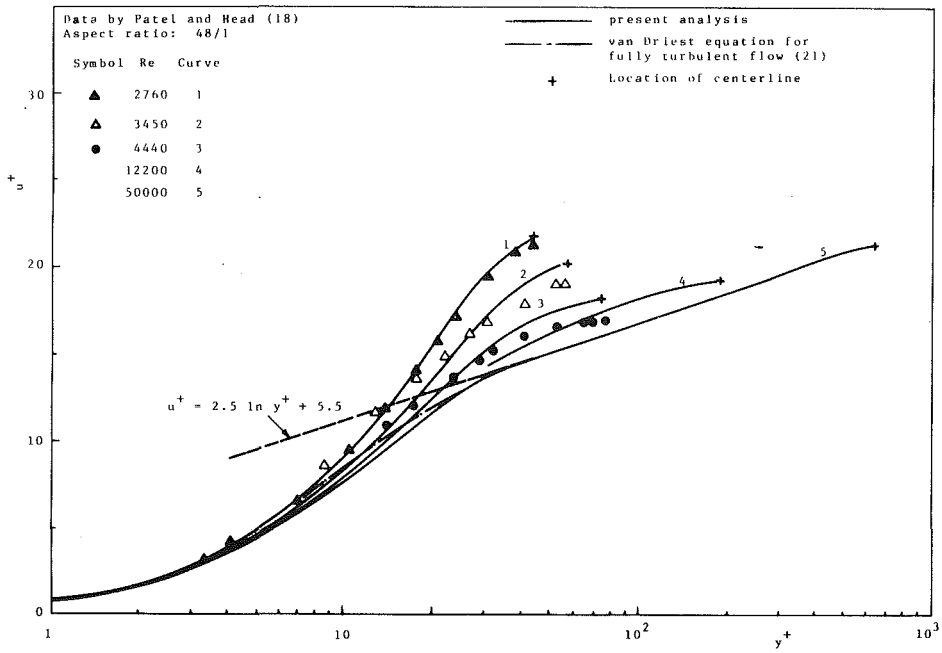


Fig. 4 Dimensionless axial velocity distribution for fully developed transitional turbulent channel flow

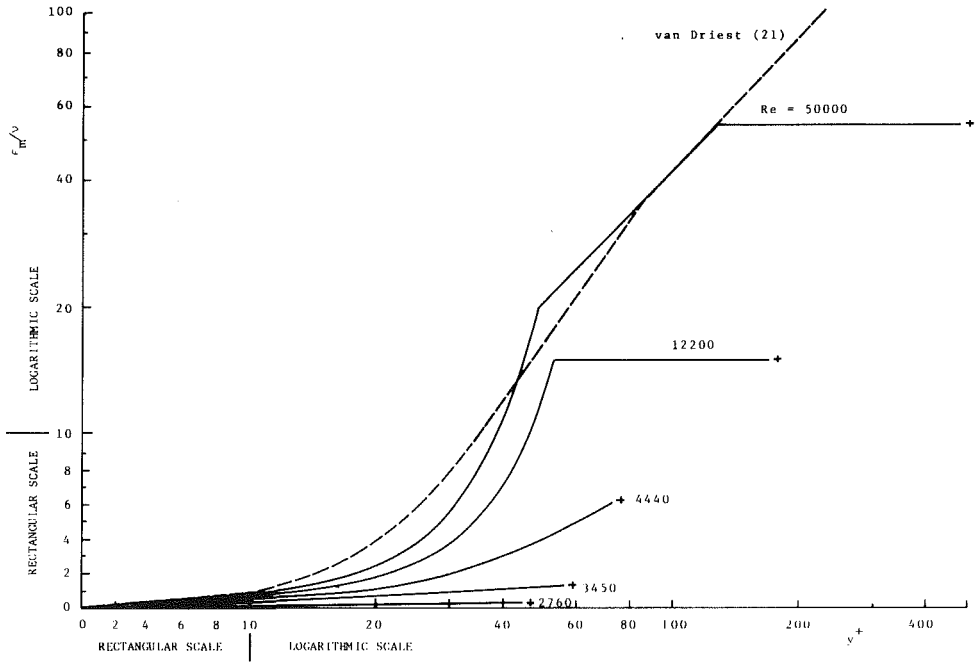


Fig. 5 Predictions for eddy viscosity for fully developed transitional turbulent flow

such conditions are shown in Fig. 1 to be in basic agreement with the surface-renewal-based correlation for fully turbulent conditions (17)

$$\frac{u_\tau}{\sqrt{\bar{s}\nu}} = 16 \tag{32}$$

This correlation for fully turbulent conditions together with the fact that $\bar{s}$ must be equal to zero for Re less than the critical transition Reynolds number Re_c provides a base for developing a reasonable estimate for $\bar{s}$ in the transitional turbulent region. To establish a working expression for $\bar{s}$, equation (32) is shown in terms of s^* and Re in Fig. 2, with f approximated by the following correlation for fully turbulent conditions:

$$f = 0.046\ \mathrm{Re}^{-0.2} \tag{33}$$

As a first approximation, s^* is assumed to take form

$$\sqrt{s^*} = A + B\ ln\ \mathrm{Re} \tag{34}$$

with A and B selected such that $s^* = 0$ at $\mathrm{Re} = \mathrm{Re}_c$, and s^* interfaces tangentially with equation (32). For the value of Re_c reported by Patel and Head [18], equation (34) reduces to

$$\sqrt{s^*} = -108 + 13.7\ ln\ \mathrm{Re} \quad \mathrm{Re}_c \ll \mathrm{Re} \ll \mathrm{Re}_1 \tag{35}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where $Re_c = 2600$ and $Re_1 = 7900$. For Re greater than or equal to Re_1, $\bar{s}$ is approximated by equation (32); that is,

$$\sqrt{s^*} = \frac{Re}{32}\sqrt{\frac{f}{2}} \quad Re_1 < Re \tag{36}$$

With s^* specified by equations (35) and (36), predictions can now be obtained for u^+ and f for transitional turbulent flow.

Results and Discussion

Predictions for Fanning friction factor f obtained on the basis of this analysis [19] are shown in Fig. 3 for low, moderate, and high Re transitional turbulent conditions. The model predictions for f are in exceptional agreement with the experimental data of Patel and Head [18] for Re less than about 6000, but are in better agreement with the empirical correlation of Beavers et al. [20] in the moderate and high Re transitional turbulent region. Predictions obtained on the basis of the two-equation $K-\epsilon$ approach of Jones and Launder [5] are also shown. As Re falls toward the critical value, the predictions for f approach a well-known equation for laminar flow between parallel plates

$$f = \frac{24}{Re} \tag{37}$$

and as Re increases toward 30,000, the model predictions approach fully turbulent correlation given by equation (33).

The predictions for dimensionless mean velocity distribution u^+ are compared with the data by Patel and Head [18] in Fig. 4 and are in basic agreement with the data for low Re transitional turbulent conditions. As Re decreases toward Re_c, the predictions for u^+ reduce to the appropriate equation for laminar flow

$$u^+ = w^+\left[\frac{1}{4} - \left(\frac{Y^+}{w^+}\right)^2\right] \tag{38}$$

or

$$\bar{u} = -\frac{1}{2\mu}\frac{d\bar{P}}{dx}\left[\left(\frac{w}{2}\right)^2 - y^2\right] \tag{39}$$

On the other hand, as Re increases toward 30,000, the predictions in the inner region approach the surface-renewal-based inner law for u^+ developed for fully turbulent conditions (17),

$$u^+ = 16\left[1 - \exp\left(-\frac{y^+}{16}\right)\right] \qquad 0 \leqslant y^+ \leqslant 46.8 \tag{40a}$$

$$u^+ = \frac{1}{\kappa} ln\, y^+ + C_1 \qquad 46.8 \leqslant y^+ \lesssim 0.2\,\delta^+ \tag{40b}$$

where $\kappa = 0.4$ and $C_1 = 5.5$. For purpose of comparison, the van Driest [22] equation for fully turbulent conditions is also shown in Fig. 4.

Referring to the experimental data for $\bar{s}$ shown in Fig. 1, the correlation given by equation (32) was developed for fully turbulent conditions (Re $\tilde{>}$ 30,000). The first-order approximation for $\bar{s}$ used in this work for transitional turbulent conditions is shown for comparison by combining equations (35) and (36) with the model predictions for f. This practical correlation for transitional turbulent conditions is compatible with the available data for $\bar{s}$.

Because the surface-renewal approach to modeling wall turbulence does not involve the use of inputs for eddy viscosity, except in the overlying turbulent core, this modeling approach can be utilized to develop predictions for ϵ_m in the wall region. The defining equation for ϵ_m is

$$\frac{\epsilon_m}{\nu} = \frac{1 - y^+/w^+2}{du^+/dy^+} - 1 \tag{41}$$

Introducing equation (19), ϵ_m is given by

$$\frac{\epsilon_m}{\nu} = \left(1 - \frac{y^+}{w^+/2}\right)\frac{\sinh(\sqrt{s^*}/2)}{\sinh[\sqrt{s^*}(1 - y^+/w^+/2)]} - 1 \tag{42}$$

This equation is shown in Fig. 5 for several values of Reynolds number in the transitional turbulent zone. To complete the picture, the correlations for ϵ_m within the outer region, equations (23) and (24), are also shown. As Re decreases from a fully turbulent value of the order of 30,000 to the laminar/turbulent transition value, the predictions for ϵ_m are observed to range from values near the van Driest correlation for fully turbulent flow to zero. As Re decreases in the transition region, the increasing extent of the wall region becomes more and more apparent.

Conclusion

Predictions for u^+ and f obtained on the basis of the surface-renewal model have been found to be in good agreement with experimental data for fully developed transitional turbulent channel flow. For low Re transitional turbulent conditions, for which case the wall region extends across the entire flow field, this model of the turbulent burst phenomenon requires the specification of only one empirical input, s^*. In contrast, the $K-\epsilon$ transport equation approach of Jones and Launder [5] necessitates the use of five empirical constants. For higher values of Re, model closure has been accomplished by the specification of s^* and by matching the wall model with a simple eddy viscosity representation of the turbulent core. With regard to the model predictions for ϵ_m, no published experimental data is available for comparison, except for limiting conditions associated with laminar flow and fully turbulent flow. The model predictions for ϵ_m falls to zero as Re approaches the laminar/turbulent transition point, and is in good agreement with the van Driest equation for fully turbulent conditions.

The input for the key modeling parameter $\bar{s}$ is compatible with experimental data for moderate to high Re transitional turbulent conditions and fully turbulent conditions, and is equal to zero for laminar conditions. The appropriateness of this input for $\bar{s}$ is further reinforced by the model predictions for u^+ and f throughout the transitional turbulent zone. However, as experimental measurements become available for turbulence characteristics in the wall region for low Re transitional turbulent conditions, refinements will undoubtedly be made in the input for $\bar{s}$.

As pointed out by Reynolds (21), the fundamental advantage of the surface-renewal approach is its relation to a specific and not unrealistic picture of events in the viscous sublayer. Because of its sound physical basis, this modeling approach provides a means for generalization to account for other complicating factors associated with transitional turbulent flow, such as curvature and heat transfer.

References

1 Pai, S. E., "On Turbulent Flow Between Parallel Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 20, 1953, pp. 109–114.

2 Wilson, N. W., and Azad, R. S., "A Continuous Prediction Method For Fully Developed Laminar, Transitional, and Turbulent Flow in Pipes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, 1975, pp. 51–54.

3 Gill, N. W., and Scher, M., "A Modification of the Momentum Transport Hypothesis," *AIChE J.*, Vol. 7, 1961, pp. 61–63.

4 Jones, W. P., and Launder, B. E., "The Calculation of Low Reynolds Number Phenomena with a Two-Equation Model of Turbulence," *Int. J. Heat Mass Transfer*, Vol. 16, 1973, pp. 1119–1129.

5 Jones, W. P., and Launder, B. E., "The Prediction of Laminarization with a Two-Equation Model of Turbulence," *Int. J. Heat Mass Transfer*, Vol. 15, 1972, pp. 301–314.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

6 Harlow, F. H., and Nakayama, P. I., "Transport of Turbulence Energy Decay Rate," Los Alamos Sci. Lab Rept., 3854, 1968.

7 Lawn, C. J., "Turbulent Heat Transfer at Low Reynolds Numbers," ASME *Journal of Heat Transfer,* Vol. 91, 1969, pp. 532–536.

8 Spalding, D. B., "Turbulence Models for Heat Transfer," *Proc. Sixth Int. Heat Transfer Conference,* Vol. 6, 1978, p. 33.

9 Kays, W. M., and Crawford, M. E., *Convection Heat and Mass Transfer,* McGraw Hill, New York, 1980.

10 Thomas, L. C., and Kakarala, C. R., "A Unified Model for Turbulent and Laminar Momentum Transfer: Channel Flow," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 8–12.

11 Einstein, H. A., and Li, H., "The Viscous Sublayer Along a Smooth Boundary," *ASCE J. Engineering Mechanics Division,* Vol. 82, 1956, pp. 1–27.

12 Hanratty, T. J., "Turbulence Exchange of Mass and Momentum With a Boundary," *AIChE J.,* Vol. 2, 1956, pp. 359–362.

13 Thomas, L. C., "Temperature Profiles for Liquid Metals and Moderate Prandtl Number Fluids," ASME *Journal of Heat Transfer,* Vol. 92, 1970, pp. 565–567.

14 Thomas, L. C., "A Turbulent Burst Model of Wall Turbulence for Two-Dimensional Turbulent Boundary Layer Flow," *Int. J. Heat and Mass Transfer,* in press, 1982.

15 Davies, J. T., and Lozano, F. J., "Turbulence and Surface Renewal at the Clean Surface of a Stirred Vessel," Symposium on Fundamental Research in Fluid Mechanics, AIChE National Meeting, New Orleans, La., 1981.

16 Danckwerts, P. V., "Significance of Liquid—Film Coefficients in Gas Absorption," *I & EC,* Vol. 43, 1951, pp. 1460–1467.

17 Thomas, L. C., "Turbulent Burst Phenomenon: Inner Laws for u^+ and T^+," NATO Advanced Study Institute in Turbulence, Istanbul, 1979.

18 Patel, V. C., and Head, M. R., "Some Observations on Skin Friction and Velocity Profiles in Fully Developed Pipe and Channel Flow," *J. Fluid Mech.,* Vol. 38, 1969, pp. 181–201.

19 Kadry, H. M., "A Model for Transitional Turbulent Convection Heat Transfer," M. S. Thesis, University of Petroleum and Minerals, Dhahran, Saudi Arabia, Feb. 1981.

20 Beavers, G. S., Sparrow, E. M., and Lloyd, J. R., "Low Reynolds Number Turbulent Flow in Large Aspect Ratio Rectangular Ducts," *J. Basic Eng.,* Vol. 93, 1971, pp. 296–299.

21 Reynolds, A. J., "The Prediction of Turbulent Prandtl and Schmidt Numbers," *Int. J. Heat and Mass Transfer,* Vol. 18, 1975, pp. 1055–1068.

22 van Driest, E. R., "On Turbulent Flow Near a Wall," *J. Aero. Sci.,* Vol. 23, 1956, pp. 1007–1017.

23 Thomas, L. C., and Greene, H. L., "An Experimental and Theoretical Study of the Viscous Sublayer for Turbulent Tube Flow," Proc. Symposium on Turbulence in Liquids, The University of Missouri, Rolla, Mo., 1973.

24 Meek, R. L., and Baer, A. P., "The Periodic Viscous Sublayer in Turbulent Flow," *AIChE J.,* Vol. 16, 1970, p. 841.

25 Rao, K. N., Narasimha, R., and Badri Narayanan, M. A., Indian Inst. of Science Report, 69FM8, 1969.

26 Kim, H. T., Kline, S. J., and Reynolds, W. C., Stanford University Report MD20, Stanford, Calif., 1969.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

S. S. Chen
Mem. ASME

J. A. Jendrzejczyk
Mem. ASME

Components Technology Division,
Argonne National Laboratory,
9700 South Cass Avenue,
Argonne, Ill. 60439

Experiment and Analysis of Instability of Tube Rows Subject to Liquid Crossflow

A tube array subjected to crossflow may become unstable by either one or both of the two basic mechanisms: velocity mechanism and displacement mechanism. An experimental and analytical study of tube rows in liquid crossflow is performed to verify the existence and transition between the two mechanisms at the intermediate values of mass-damping parameter. Experimental data and analytical results are found to be in good agreement.

1 Introduction

Fluid flowing across a tube array can cause structural instability. These problems have become more important with the development of more efficient and larger equipment. Reviews of this subject are available [e.g., 1-3].

In the past, instability of a tube array in crossflow was thought to be attributed to fluid-elastic force only. A general theory of stability of tube arrays subjected to crossflow was developed recently [4, 5]. It is shown that dynamic instability may be caused by two basic mechanisms:

1. Velocity mechanism: The dominant fluid force is proportional to the velocity of tubes; this force is called the fluid-damping force. Depending on the reduced flow velocity, the dominant fluid-damping force may act as an energy-dissipation mechanism or an excitation mechanism for tube oscillations. When it acts as an excitation mechanism, the system damping is reduced. Once the modal damping becomes negative, the tubes lose stability. This type of instability is called fluid-damping-controlled instability.

2. Displacement mechanism: The dominant fluid force is proportional to the displacement of tubes; this force is called the fluid-elastic force. The dominant force may affect the natural frequency as well as modal damping. As the flow velocity increases, the fluid-elastic force may reduce the modal damping. When the modal damping becomes negative because of the fluid-elastic force, the tubes become unstable by the displacement mechanism; this type of instability is called fluid-elastic, stiffness-controlled instability.

In most cases, either the velocity mechanism or the displacement mechanism will be dominant. For a tube row with a pitch-to-diameter ratio of 1.33, analytical results show that instability is attributed to the velocity mechanism for small mass-damping parameter (δ_m) and to the displacement mechanism for large δ_m.

This paper presents an experimental and analytical investigation of the instability phenomena of tube rows in liquid crossflow to verify the existence and transition between the two mechanisms. Two series of tests were performed to establish the critical flow velocity as a function of the mass-damping parameter and to verify the discontinuity of the critical flow velocity at a certain value of δ_m. Experimental data and analytical results are found to be in good agreement. Therefore, the two basic mechanisms are established both analytically and experimentally.

2 Test Section and Test Procedures

The experiments are performed in the rectangular flow channel, 10.7 cm (4.2 in.) wide and 25.9 cm (10.2 in.) high, of a test chamber connected to a water loop with a maximum flow rate of 0.052 m^3/s (700 gpm). The rectangular flow channel is formed by four internal walls which serve to separate the internal flow from the almost stagnant water contained in the surrounding cross section of the external 30 cm (12 in.) diameter pressure shell. A section of this shell is fitted with two vertical pipe branches which accommodate a test assembly. The detail of the test chamber is the same as that used for earlier experiments for tube arrays in crossflow [6].

The arrangement of the tube row is shown in Fig. 1. Tube outside diameter (D) is 1.59 cm (5/8 in.), and tube wall thickness is 0.079 cm (1/32 in.). The transverse pitch-to-diameter ratio is 1.35.

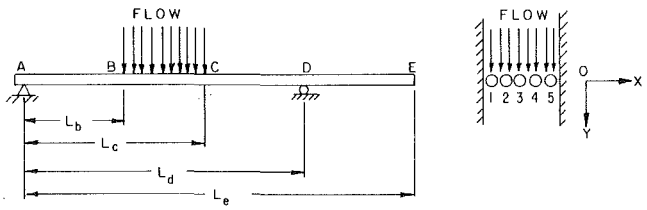


Fig. 1 Schematic of the tube row

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October, 1981; final revision, April, 1982. Paper No. 82-WA/APM-18.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

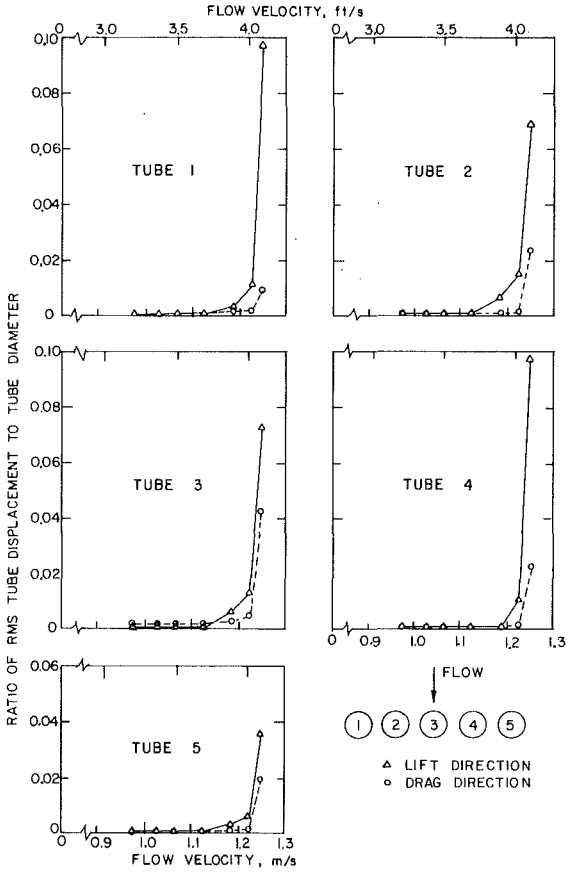


Fig. 2 Tube displacement as a function of flow velocity in test *a*.1 at 23°C

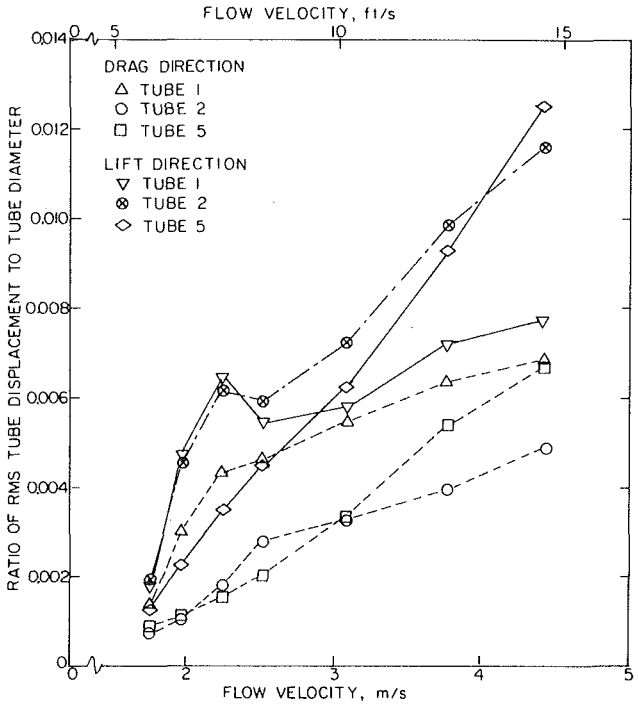


Fig. 3 Tube displacement as a function of flow velocity in test *a*.1 at 8.9°C

Each tube element is suspended as a simply supported beam on two *O*-rings (*A*, *D*) mounted 91.4 cm (36 in.) apart, permitting lateral vibration with equal flexibility in all directions (see Fig. 1). The *O*-rings are seated in the compression plates which are attached to the vertical pipe branches. The middle portion of the tubes *BC* (25.9 cm) is subjected to flow, *AD* is in still water, and *DE* is in air. Two sets of tubes with different lengths of overhang (*DE*) are tested: Case *a*, tubes with an overhang of 35.6 cm; and Case *b*, tubes with an overhang of 5.1 cm.

Three sequences of tests are performed for Case *a*: (1) all tubes are in tune; (2) a weight of 66.3 grams is attached to the free end of tubes 2 and 4; and (3) a weight of 66.3 grams is attached to tube 3 only. These tests will be referred to as Tests *a*.1, *a*.2, and *a*.3.

Two sequences of tests are performed for Case *b*: (1) all tubes are in tune and (2) tubes 1 and 2, and tubes 4 and 5 are connected together using a rubber plug, which is pushed into the gap between the overhang ends of these tubes. These tests are referred as Tests *b*.1 and *b*.2.

The instrumentation for the tubes consists of two accelerometers mounted at the middle of a tube with orientation to be sensitive in the drag and lift directions. The accelerometers are installed by means of a small mounting block attached to the inside of the tubes and the leads brought to the outside. The tube accelerometer outputs are amplified by charge amplifiers and coupled to double-integration circuits and an analog tape recorder. Displacement signals obtained from the double integrators are then analyzed using the FFT analyzer.

Tests for each case are made in air, still fluid, and flowing fluid.

2.1 Air Environment. Testing in air is performed to determine the natural frequency and damping of each tube. The tube is excited by plucking it at the free end. The transient response of the tube is recorded. Then, the damping is obtained from the log decrement of the displacement trace, and the natural frequency is determined from the power spectrum of the displacement. Each instrumented tube is tested in two orthogonal directions.

2.2 Still Fluid. Uncoupled natural frequency and modal damping ratio are determined using the same procedure as in air. In this test, all the tubes surrounding the tubes being tested are held at the overhung end by hands to eliminate responses of those tubes.

2.3 Flowing Fluid. The flow velocity is increased at small intervals. At each flow velocity, the displacement signals in the lift and drag directions are recorded on a magnetic tape for several minutes for subsequent analysis. The test is terminated when tube impact occurs.

Tube damping depends on water temperature; this is attributed to *O*-rings, whose characteristics are a function of temperature [6]. Therefore, tube damping can be controlled by controlling water temperature. In each test sequence, the tube row is tested under several different temperatures: in-air test at room temperature, and in-fluid and in-flow tests at several loop temperatures with one of them being equal to room temperature.

Nomenclature

D = diameter of tubes
f = natural frequency (Hz) in air
L_e = tube length
m = mass per unit length of tubes
U = gap flow velocity
δ_m = mass-damping parameter $\left(=\dfrac{2\pi\zeta m}{\rho D^2}\right)$
ζ = averaged modal damping ratio in air
ρ = fluid density

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

3 Experimental Data

3.1 Natural Frequency and Damping. Tube natural frequency can be determined from the frequency spectra and damping from the displacement-time traces. Natural frequency and damping are measured for each tube in the lift and drag directions. Because of length limitation, detailed results are not included in the paper.

Both natural frequency and damping in still water increase with decreasing temperature. The effect on natural frequency is not very significant. But damping increases significantly at lower temperatures.

3.2 Tube Response in Flow. Figures 2 and 3 show typical tube displacements as a function of flow velocity. The movements of the tubes are strongly dependent on flow velocity. At low flow velocities, the motion is small. As the flow velocity increases, tube displacement increases. When the flow velocity exceeds a certain value, rapid increase in vibration occurs; the threshold flow velocity is called the critical flow velocity. In the flow-velocity range tested, tube displacement in the lift direction is larger than in the drag direction.

Tube responses depend on damping, which is a function of water temperature. Figures 2 and 3 show significant differences in responses at two different temperatures. At 23°C, the tubes become unstable at 1.22 m/s, but at 8.9°C, up to about 4.5 m/s, the tubes are still stable.

Tube responses also depend on frequency variation among the constituent tubes. When the tubes are in tune, the tubes become unstable at a lower flow velocity. When the tubes are out of tune, the critical flow velocity increases. For example, the critical flow velocity for Test *b*.2 is larger than that of Test *b*.1.

Tube responses exhibit nonlinear characteristics. In several tests, the flow velocity is increased at several intervals until instability occurs. Then the flow velocity is gradually reduced. Tube displacements with decreasing flow rate are different from those of increasing flow rate. The tubes are still in unstable motion, even though the flow velocity is reduced to those lower than the critical value. With further decrease in flow velocity, the tubes eventually regain stability, but the flow velocity to regain stability is smaller than the critical value.

Additional tests have also been made to understand the significance of this nonlinear phenomenon using Test *a*.1. First, the flow velocity is increased to a little above the critical value. At this flow velocity, the tubes are excited to large oscillations (instability). Then, the flow velocity is decreased to a little less than the critical value; the tubes are still performing large oscillations. As soon as the tubes are held by hand at the overhung portion, the tubes regain stability. The tubes remain stable after the hand is removed. Therefore, the nonlinear characteristics possess little practical significance and the critical flow velocity should be established using the case of increasing flow rate.

Typical frequency spectra of tube displacements are given in Figs. 4 and 5. Tube response frequencies are contained within the frequency band of coupled modes [7]. In the subcritical flow-velocity ranges, tube response frequencies increase with flow velocity; this can be seen clearly from Fig. 5. Once the tubes become unstable, the response frequency is associated with a particular mode.

Figure 4 shows that there is a shift in frequency between 1.36 and 1.40 m/s. At 1.36 m/s, the tubes are stable, but at 1.40 m/s, the tubes become unstable. When the flow velocity is gradually decreased, the response frequency also gradually decreases and eventually returns to the value before instability occurs. Since at instability, the tubes respond at a particular well-defined frequency, the critical flow velocity can also be

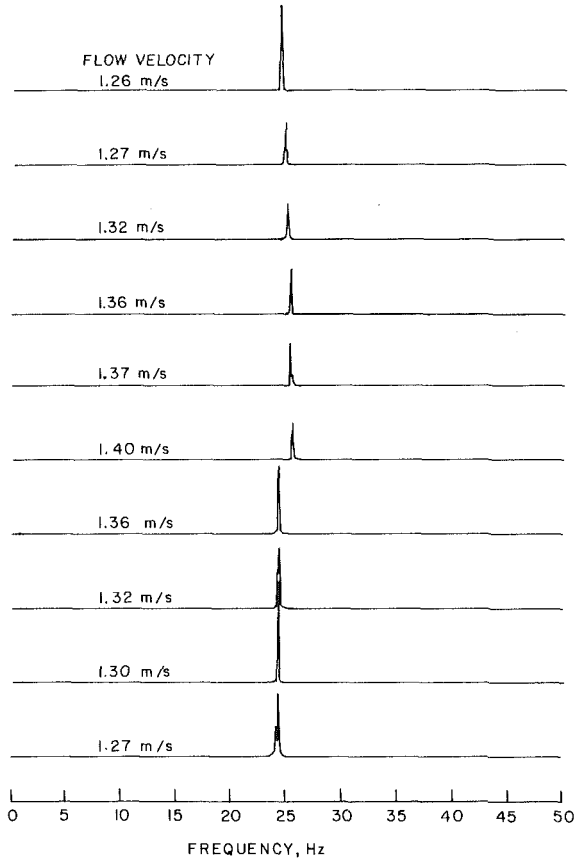


Fig. 4 Frequency spectra of tube displacements of tube 3 in the lift direction in test *a*.1 at 16.7°C; sequence of tests is from bottom to top

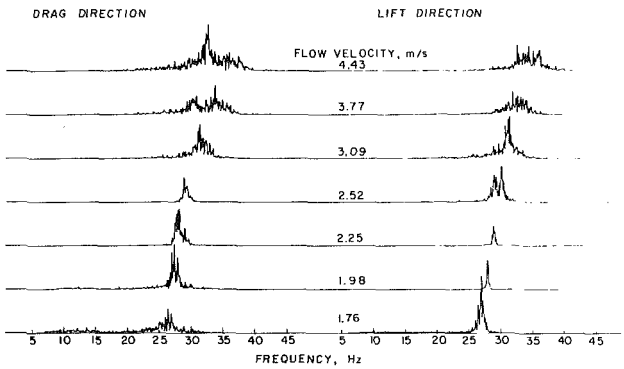


Fig. 5 Frequency spectra of tube displacements of tube 2 in test *a*.1 at 8.9°C

established by examining the frequency spectra in conjunction with the tube displacements.

Figure 6 shows the instability modes of Tests *a*.1, *a*.2, and *a*.3. In Test *a*.1, all tubes are in tune; the amplitudes of the tubes are about the same except tube 5. In Test *a*.2, tubes 2 and 4 have a lower natural frequency; the displacements of these tubes are much larger than the others. Apparently, tubes 2 and 4 first become unstable. In Test *a*.3, tube 3 has a lower natural frequency, but the displacement of tube 3 is smaller than the others. In this case, tubes 1 and 2, and 4 and 5, are in tune. The instability is apparently associated with those tubes.

3.3 The Critical Flow Velocity. One of the main objectives is to establish the critical flow velocity. It is defined as the flow velocity at which the tube displacements increase sharply. In all tests, the critical flow velocities are fairly well defined. The critical flow velocity is also characterized by a

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

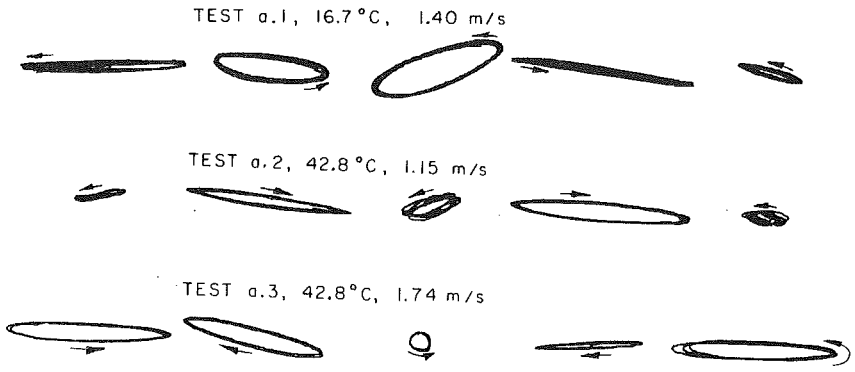


Fig. 6 Tube orbital paths of tubes at instability in tests *a*.1, *a*.2, and *a*.3

Table 1 Measured critical flow velocity and associated instability frequency

Tests	Water temperature, °C	Critical flow Velocity, m/s	Oscillation frequency (Hz) at instability
	42.8	1.07	24.3
	23.0	1.19	24.3
	16.7	1.28	24.3
a.1	11.1	1.51	25.5
	9.4	1.83	26.8
	9.1	2.00	28.0
	8.9	5.18	not measured
	42.8	1.07	24.0
a.2	23.0	1.23	24.4
	11.1	1.69	26.9
	42.8	1.07	21.4
a.3	23.0	1.23	22.0
	11.1	1.95	28.7
	43.3	1.13	25.4
b.1	23.3	1.30	25.7
	16.7	1.43	28.0
	11.1	1.65	29.7
	43.3	1.30	28.0
b.2	23.3	1.46	27.8
	16.7	1.61	28.3
	11.1	1.81	29.4

single tube-response frequency. The critical flow velocity and the oscillation frequency at instability for all tests are given in Table 1.

In Test *a*.1 at 8.9°C, the critical flow velocity is much higher than the others and the tubes oscillate violently. Measurements for the oscillation frequency and instability mode are not made to prevent damage caused by tubes impacting with one another.

4 Analytical and Experimental Results

The analytical method for predicting the general characteristics of tube arrays in crossflow was described in reference [4]. For conciseness, the analytical procedures will not be repeated in this paper. Based on the analytical model, natural frequencies and modal damping ratios of uncoupled and coupled modes for tube arrays oscillating in still fluid and flowing fluid can be calculated. The lowest critical flow velocity at which large tube oscillations occur can also be calculated.

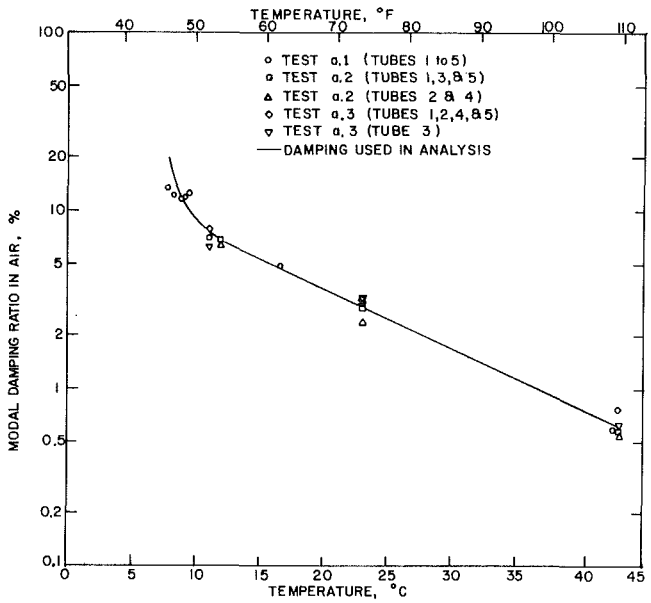


Fig. 7 Averaged modal damping ratio in air

Figure 7 shows the modal damping ratio in air calculated from the analytical model and measured damping values in still fluid. In still water, the modal damping values are different in the drag and lift directions. However, in air, their values in the two directions are about the same. In Fig. 7, the averaged values of damping in the two directions are presented. For calculations of the critical flow velocity, the solid line given in Fig. 7 is used.

Calculation of the "exact" critical flow velocity based on a row of five tubes is straightforward [4, 5]. Fluid-force coefficients used in the analysis are based on those for a tube row with a pitch-to-diameter ratio of 1.33. Computations are made for a row of five identical tubes corresponding to Tests *a*.1 and *b*.1 for different values of damping ratio; the results are presented by the solid lines given in Figs. 8 and 9. The instability modes are also given.

Dashed lines in Figs. 8 and 9 present the theoretical critical flow velocity based on constrained modes [4, 5]. For smaller values of mass-damping parameter, the dashed line corresponds to a single elastic tube in a rigid tube row, but for large values of mass-damping parameter, the dashed line

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

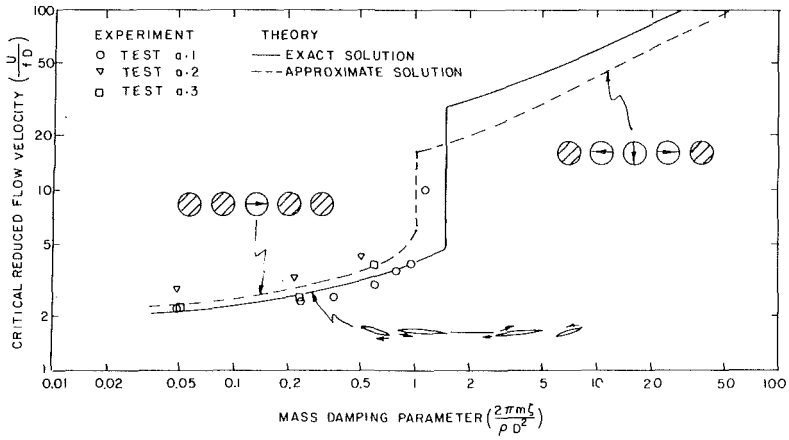


Fig. 8 Critical flow velocity for case *a*

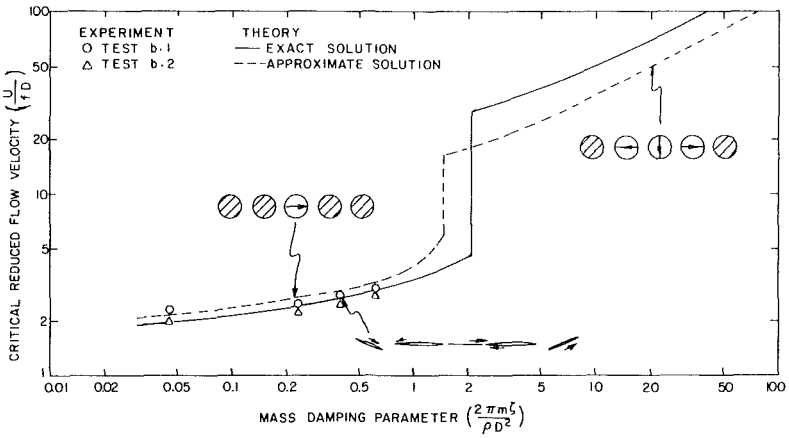


Fig. 9 Critical flow velocity for case *b*

corresponds to a group of three elastic tubes in a tube row with the middle tube moving in the drag direction and the two wing tubes moving in the lift direction.

The experimental values of the critical reduced flow velocity (U/fD) and mass-damping parameter ($2\pi m\zeta/\rho D^2$) are evaluated based on the experimental data of natural frequency (f), critical flow velocity given in Table 1, and damping values given in Fig. 7. The experimental results are also presented in Figs. 8 and 9.

The experimental and analytical results for the critical flow velocity given in Figs. 8 and 9 agree well. Note that the exact solution in Fig. 8 is corresponding to Test $a.1$ and the solid line in Fig. 9 is corresponding to Test $b.1$. Both of them agree fairly well.

At a certain value of δ_m ($=2\pi m\zeta/\rho D^2$), there is a jump of the critical flow velocity. The jump is attributed to the transition from one dominant instability mechanism to another dominant instability mechanism. For δ_m smaller than this value, the instability is of the fluid-damping-controlled type and for larger δ_m, the instability is of fluid-elastic, stiffness-controlled type. This jump is demonstrated in Test $a.1$. With a small increase in damping by decreasing water temperature less than 1 deg. the critical flow velocity increases several times. The analytical result for δ_m at which jump occurs agrees with the experimental data reasonably well.

Experimental instability modes given in Fig. 6 show that the most critical mode is the out-of-phase mode in the lift direction, in which the adjacent tubes are moving out of phase to each other. This is in agreement with the analytical instability modes given in Figs. 8 and 9. For tube rows with low δ_m, this mode is the critical instability mode.

The critical flow velocities for Test $a.1$ are lower than those for Tests $a.2$ and $a.3$, and the critical flow velocities for Test $b.1$ are lower than those for Test $b.2$. This demonstrates the effect of detuning, which tends to stabilize the tube row. Analytical results also show that detuning is beneficial in the stability of a tube array [4, 5].

Calculations also have been made to predict the coupled natural frequencies of the tube row as a function of flow velocity. The general characteristics are similar to those given in Fig. 5; i.e., the coupled natural frequencies in the subcritical flow-velocity range increase with flow velocity.

5 Conclusions

This study, for the first time, demonstrates experimentally the existence of two instability mechanisms and the jump in critical flow velocity at a certain value of the mass-damping parameter.

Based on the experimental and analytical results, the following conclusions can be made.

1. There are two basic instability mechanisms: fluid-damping-controlled instability and fluid-elastic, stiffness-controlled instability. For a tube row with a pitch-to-diameter ratio of 1.35, the former is dominant for small mass-damping parameter, and the latter is dominant for large mass-damping parameter.

2. There is a finite jump in the critical flow velocity at a certain value of mass-damping parameter associated with the transition from one instability mechanism to another. The jump phenomenon is demonstrated in experiment by varying tube damping.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

3. At low mass-damping parameter, fluid-damping force is dominant and, for a tube row, the dominant instability mode is associated with the out-of-phase mode in the lift direction.

4. Detuning of the tubes tends to increase the critical flow velocity. However, for fluid-damping-controlled instability, its effect is not very significant.

The experimental data have verified the validity of the mathematical model [4, 5]; thus, the model has established the foundation for the development of a design guide to avoid detrimental instability.

Acknowledgment

This work was performed under the sponsorship of the Office of Reactor Research and Technology, U.S. Department of Energy.

References

1 Chen, P. Y., ed., "Flow-Induced Vibration Design Guidelines," ASME, PVP, Vol. 52, 1981.

2 Blevins, R. D., *Flow-Induced Vibration*, Van Nostrand Reinhold, New York, 1977.

3 Chen, S. S., "Cross-Flow-Induced Instabilities of Circular Cylinders," *The Shock and Vibration Digest*, Vol. 12, No. 5, 1980, pp. 21–34.

4 Chen, S. S., "Instability Mechanisms and Stability Criteria of a Group of Circular Cylinders Subjected to Cross Flow; Part I: Theory," ASME Paper No. 81-DET-21, to appear in ASME *Journal of Mechanical Design*.

5 Chen, S. S., "Instability Mechanisms and Stability Criteria of a Group of Circular Cylinders Subjected to Cross Flow; Part II: Numerical Results and Discussions," ASME Paper No. 81-DET-22, to appear in ASME *Journal of Mechanical Design*.

6 Chen, S. S., Jendrzejczyk, J. A., and Lin, W. H., "Fluidelastic Instability of Rectangular Tube Arrays Subjected to Liquid Cross Flow," *Proc. of the Int. Conf. on Pressure Vessel Technology*, London, 1980, pp. 193–200.

7 Chen, S. S., "Vibrations of a Row of Circular Cylinders in a Liquid," ASME *Journal of Engineering for Industry*, Vol. 97, 1975, pp. 1212–1218.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

G. Francfort
Research Assistant.

A. Golebiewska Herrmann
Acting Associate Professor.
Mem. ASME

Division of Applied Mechanics,
Department of Mechanical Engineering,
Stanford University,
Stanford, Calif. 94305

Conservation Laws and Material Momentum in Thermoelasticity

The main goal of this paper is to construct a Lagrangian function such that not only the well-known equations of thermoelasticity, but also material conservation laws can be derived. As action variables, the position $\mathbf{x}$ *of a material particle and a scalar function* η *related to temperature are used. The material momentum for thermoelasticity is derived. Here, by contrast to the purely elastic case, the material momentum depends on a time interval rather than on an instant of time. The balance of material momentum is integrated over time to produce a relation reminiscent of the impulse-momentum equation in classical mechanics.*

1 Introduction

The development of mathematical fracture mechanics focused attention on new material conservation laws because they form a starting point for establishing path-independent integrals. These conservation laws were first derived for a static linear-elastic continuum. Numerous attempts have been made to generalize them for nonlinear, plastic, viscoelastic, and dynamic behavior.

It turns out that these new material conservation laws fit naturally into a physical framework: they are related to the (material) symmetries of the Lagrangian function of the given system which in a static elastic case reduces to the strain energy.

A natural obstacle prevents generalization to thermoelastic systems. The dissipative character of these systems precludes the generation of a Lagrangian for thermoelasticity by simple addition of new terms, involving the fields (which describe the thermal behavior) to the elastic Lagrangian. As long as we are interested in deriving the thermoelastic equations only, it is possible to overcome this obstacle by using adjoint fields as well as original ones in the Lagrangian. Such a formulation, however, is not capable of delivering the new conservation laws because the crucial quantity appearing there involves the Lagrangian itself.

The current paper is an attempt to construct a Lagrangian, from which we can derive all known equations of thermoelasticity, as well as new conservation laws.

2 General Background

The variational approach to many physical theories proved very useful over the years and will be followed here. For a system of particles, for example, we consider an action integral, where the integrand is called the Lagrangian of the system under consideration and the independent variable is the time t. In various field theories, by contrast, we deal with a Lagrangian density because the integration is performed not only over an interval of time, but over a region of space V. In other words, the actions A have the following forms

$$A = \int_{t_0}^{t_1} L(t;\mathbf{x}_1,\mathbf{x}_2 \ldots \mathbf{x}_n,\mathbf{v}_1,\mathbf{v}_2 \ldots \mathbf{v}_n)\, dt \tag{1}$$

for systems of n particles, and

$$A = \int_{t_0}^{t} \int_{V} \mathcal{L}(t,\mathbf{x};\phi,\dot{\phi},\phi_{,i})\, dt\, d^3x \tag{2}$$

for fields. Here $\mathbf{x}_i$ $(i = 1, 2, \ldots n)$ denotes the position of the ith particle, $\mathbf{v}_i$ is its velocity, and all $\mathbf{x}_k$, $\mathbf{v}_k$ are functions of t. In (2) ϕ is a field (or fields), a function of both, and time t and space coordinates $\mathbf{x}$, $\dot{\phi}$, and $\phi_{,i}$ denote partial derivatives with respect to time and x_i, respectively.

According to the stationarity principle, by varying the dependent variables ($\mathbf{x}$ in the first case, ϕ in the second), we derive the equations of motion for the particles or the field equations. In many cases we are also interested in conservation laws for the given system. They can be derived from Noether's theorem [1] and are related to the invariance of the action integral under infinitesimal transformations performed on independent and/or dependent variables.

The Lagrangian description of the continuum can be derived from (1) when the number n becomes very large and particles are densely distributed. $\mathbf{x}_n(t)$ then becomes a continuous function of a material coordinate $\mathbf{a}$ (related to the original number of a particle n) i.e., $\mathbf{x} = \mathbf{x}(\mathbf{a},t)$. Most often a displacement vector $\mathbf{u} = \mathbf{x} - \mathbf{a}$ is used for the description of a continuum rather than the position $\mathbf{x}$ itself. Then relation (1) transforms into

$$A = \int_{t_0}^{t} \int_{B} L(\mathbf{a},t;\mathbf{u},\mathbf{v},\nabla\mathbf{u})\, d^3a\, dt \tag{3}$$

where the integration is performed over the whole body B. Here $\mathbf{v}$ and $\nabla\mathbf{u}$ denote partial derivatives of $\mathbf{u}$ with respect to t and a_i $(i = 1,2,3)$. The usually accepted equations of an

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1982. Paper No. 82-WA/APM-26. Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

elastic continuum are not, however, generally obtained from stationariness of the action integral (3), but rather by extending well-known principles for the discrete systems of particles to a continuum. In other words, the introduction of the new independent variables **a** was not really exploited until 1951 when Eshelby introduced the notion of a force on an elastic singularity [2], related to translations in **a** space. In the absence of defects, the relation obtained by Eshelby becomes a statement of conservation of field (or material) momentum.

New conservation laws, associated with the properties of material space, have been derived by Günther [3] and Knowles and Sternberg [4]. These conservation laws are closely linked to the path-independent integrals of defect (especially fracture) mechanics. A similar situation arises in particle mechanics where the conservation of linear momentum becomes a balance of linear momentum (or equation of motion) if external forces are present. The first of such path-independent quantities for a crack was established in 1968 by Rice [5] and extended by Budiansky and Rice [6]. Eshelby's approach and most others are essentially based on the formulation of the type represented by equation (3), e.g., they use the displacement vector as a basic field quantity. We are now confronted with a very peculiar and unusual (from the field-theroretical point of view) situation. Namely, conservation laws, in general, are related to the transformation on either dependent or independent variables as mentioned before. The equations of continuum mechanics, however, represent balance of linear and angular momentum, which in turn follow from transformations of physical coordinates. But the physical coordinates do not directly enter the formulation either as dependent or as independent variables.

To take full advantage of the Lagrangian field-theoretical formalism, we have to adopt **x** as the dependent variable (instead of **u**) in (3), or alternatively (if we want to achieve formal resemblance to field theory) we might use a formulation based on form (2), where **a** would be playing the role of the field ϕ. This has been done by Golebiewska Herrmann [7]. In this approach two basic quantities were introduced in a natural fashion into the development, namely the physical momentum tensor, i.e., the Piola-Kirchhoff stress tensor and the material momentum tensor related to Eshelby's tensor. The latter was given as

$$b_{ik} = \frac{\partial L}{\partial x_{j,k}} x_{j,i} - L\delta_{ik} \tag{4}$$

in the representation using $\mathbf{a},t$ as independent variables.

As it is seen, the Lagrangian density itself enters the expression for the material momentum tensor b_{ik}. It is therefore of prime importance to be able to produce a suitable expression for the Lagrangian. In this paper we construct material momenta for linear, dynamic, fully coupled thermoelasticity. Thus our first concern is the generation of an appropriate Lagrangian density.

3 A Short Review of Variational Formulation in Thermoelasticity

The equations of linear thermoelasticity are

$$-\sigma_{ij,j}(\mathbf{a},t) - f_i(\mathbf{a},t) + \rho(\mathbf{a})\ddot{u}_i(\mathbf{a},t) = 0 \tag{5a}$$

$$-q_{i,i}(\mathbf{a},t) - r(\mathbf{a},t) + \dot{s}(\mathbf{a},t) = 0 \tag{5b}$$

where the u_i's are the components of the displacement vector, usually defined as,

$$u_i(\mathbf{a},t) = x_i(\mathbf{a},t) - a_i$$

where $\rho(\mathbf{a})$ is the density of the medium, f_i represents body forces, σ_{ij} is the Cauchy stress tensor, and r and q_i are the heat source and the heat flux, respectively. Finally, s denotes the entropy of the system. The aforementioned quantities satisfy certain relations, namely

$$\begin{aligned} \sigma_{ij}(\mathbf{a},t) &= c_{ijkp}(\mathbf{a})\, u_{k,p}(\mathbf{a},t) - \beta_{ij}(\mathbf{a})\tau(\mathbf{a},t) \\ q_i(\mathbf{a},t) &= \lambda_{ij}(\mathbf{a})\, \tau_{,j}(\mathbf{a},t) \\ s(\mathbf{a},t) &= \beta(\mathbf{a})\, \tau(\mathbf{a},t) + \beta_{ij}(\mathbf{a})\, u_{i,j}(\mathbf{a},t) \end{aligned} \tag{6}$$

where we introduced the temperature increment $\tau(\mathbf{a},t)$ and where c_{ijkp}, β_{ij}, λ_{ij}, β are the elastic and thermoelastic coefficients. In most cases these coefficients are constant.

A characteristic feature of these equations resides in the coupling of the equations of motion (5*a*) with a differentiated form of the second law of thermodynamics (5*b*). The latter equation induces dissipation and the whole system becomes nonconservative.

The basic purpose of a variational formulation of the problem will always be the generation of the complete set of equations of Euler's equations associated with a given Lagrangian density. The dissipative character of the thermoelastic system is of essential importance and precludes the possibility of constructing the Lagrangian density by simple modification of its elastic analog (e.g., by adding additional terms).

To the authors' knowledge two methods have been suggested to overcome this difficulty. The first one is based on the introduction of adjoint fields that render the system conservative. As such, it applies to any type of dissipative system. Implementations can be found in Morse and Feshback [8] for the heat equation and in Herrmann and Tasi [9] for thermoelasticity. It is essential to note that this method generates a set of equations for the adjoint fields as well. In the case of the heat equation, the adjoint field is a solution of the backward heat equation. As such, it blows up exponentially fast after a finite time for any kind of nonzero initial data. Of course, this last feature is irrelevant as long as one's goal is simply to generate the heat equation from a Lagrangian, but it forsakes any hope to obtain a meaningful material momentum, since such a momentum always involves $\mathcal{L}$, thereby coupling field and adjoint fields.

The second method has been specifically designed for thermoelasticity by Biot [10]. A field **s**, called the entropy displacement field, is defined through

$$div\, \mathbf{s} = -s \tag{7}$$

where s is the entropy of the system (as mentioned before).

Recently Rafalski [11] has suggested a Lagrangian density based on convolution in time. This feature alleviates the obstacle created by the dissipation and avoids the trap of adjoint fields. Roughly speaking, the loss of energy of the system is compensated, through the convolution in time, by the gain of energy of that same system run backward. Formally, this is done by introducing symmetrized products of two quantities at two different times, as shown in the next section. By contrast to the method of adjoint fields, the system is run backward only as long as it is meaningful. Note, however, that Rafalski's formulation still uses entropy displacements.

The few papers mentioned in this section offer, in our opinion, a representative sample of the variational techniques currently used in thermoelasticity in order to derive the thermoelastic equations. In the next section we will construct a Lagrangian density where neither adjoint fields nor entropy displacements are considered.

Before we proceed we should mention some contributions related to new conservation laws for thermoelasticity. Gurtin [12] derived a path-independent integral for the static case. Herrmann [13] found a generalization of Eshelby's tensor for thermoelasticity. However, great difficulties arise in handling the time dependence. The material momenta produced in [13]

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

are associated with the Laplace transform in time of the original equations, and their connection with the original material momenta has not been established. The author suggested that whenever the time operator occurred, the multiplication had to be carried out as a convolution. In the work mentioned in the foregoing, the entropy displacement s has been used. (The practicality of the notion of s is questionable since it involves the solving of equation (7) explicitly.)

4 A New Lagrangian for Thermoelasticity

To our knowledge, Gurtin [14] is the first to have considered nonlocal Lagrangians in the case of elasticity. As it was previously mentioned, Rafalski then proceeded to apply this concept to thermoelasticity in [11]. In our formulation we keep this essential feature, but modify the fields previously introduced.

If f and g are two functions of time, we define

$$[f,g]_{t_0}^{+} = f(t)\, g(t_0 - t) + f(t_0 - t)\, g(t)$$

$$[f,g]_{t_0}^{-} = f(t)\, g(t_0 - t) - f(t_0 - t)\, g(t)$$

Let t_0 be a strictly positive real number. If $\{x_i\}$ $(i = 1,2,3)$ and η are functions of $\mathbf{a}$ and t, we define $\mathfrak{L}(\mathbf{a},t,t_0; x_i,\dot{x}_i,x_{i,j},\eta,\dot{\eta},\eta_{,i},\dot{\eta}_{,i})$ to be

$$-\tfrac{1}{4}\, c_{ijkp}(\mathbf{a})[x_{k,p} - \delta_{kp}, x_{i,j} - \delta_{ij}]_{t_0}^{+} + \tfrac{1}{2}\, \beta_{ij}(\mathbf{a})[\dot{\eta}, x_{i,j} - \delta_{ij}]_{t_0}^{+} + \tfrac{1}{4}\, \beta(\mathbf{a})[\dot{\eta},\dot{\eta}]_{t_0}^{+} - \tfrac{1}{4}\, \rho(\mathbf{a})[\dot{x}_i,\dot{x}_i]_{t_0}^{+} + \tfrac{1}{4}\, \lambda_{ij}(\mathbf{a})[\dot{\eta}_{,j},\dot{\eta}_{,i}]_{t_0}^{+} + \tfrac{1}{2}\, [f_i,x_i]_{t_0}^{+} - \tfrac{1}{2}\, [r,\eta]_{t_0}^{+} \tag{8}$$

where f_i and r are given functions of $\mathbf{a}$ and t and all coefficients entering (8) are assumed to be smooth, real, and symmetric and satisfy the usual hypotheses of strong ellipticity which guarantee that the equations of thermoelasticity are well posed (see equation (5)). The newly introduced field η is related to the temperature τ by the relation $\tau = \dot{\eta}$.

The reader has undoubtedly noticed the nonlocality in time of $\mathfrak{L}(\mathbf{a},t,t_0;x_i,\dot{x}_i,x_{i,j},\eta,\dot{\eta},\eta_{,i},\dot{\eta}_{,i})$. Time t_0 appears explicitly in $\mathfrak{L}$, not only in the limit of the integral. To underline this fact we will use from now on the notation $\mathfrak{L}_{t_0}$ instead of $\mathfrak{L}$. This time nonlocality will turn out to be essential for our purpose.

As a first test of the adequacy of our Lagrangian density, we should be able to easily recover the classical Lagrangian density of linear elasticity. This is achieved by canceling all terms where η is present and by collapsing t with $t_0 - t$. Then $[\dot{x}_i,\dot{x}_i]_{t_0}^{+}$ becomes $-2\dot{x}_i(\mathbf{a},t)\, \dot{x}_i(\mathbf{a},t)$ since $\partial/\partial t = -\partial/\partial(t_0 - t)$ and $\mathfrak{L}t_0$ becomes

$$-\tfrac{1}{2}\, c_{ijkp}(\mathbf{a})(x_{k,p} - \delta_{kp})(\mathbf{a},t)\,(x_{i,j} - \delta_{ij})(\mathbf{a},t) + \tfrac{1}{2}\, \rho(\mathbf{a})\, \dot{x}_i(\mathbf{a},t)\, \dot{x}_i(\mathbf{a},t) + f_i(\mathbf{a})\, x_i(\mathbf{a},t) \tag{9}$$

as announced.

We now postulate the *stationarity of A* at the field point (x_i,η). We perform the classical steps leading to Euler's equations, taking full account of the symmetry of the coefficients and of the convolution in time. We finally obtain the following set of equations for $0 \le t \le t_0$:

$$-\frac{\partial}{\partial a_j}\{c_{ijkp}(\mathbf{a})[x_{k,p}(a,t) - \delta_{kp}] - \beta_{ij}(\mathbf{a})\, \dot{\eta}(\mathbf{a},t)\} + \rho(\mathbf{a})\, \ddot{x}_i(\mathbf{a},t) - f_i(\mathbf{a},t) = 0$$

$$-\beta(\mathbf{a})\, \ddot{\eta}(\mathbf{a},t) - \beta_{ij}(\mathbf{a})\, \dot{x}_{i,j}(\mathbf{a},t) + \frac{\partial}{\partial a_j}[\lambda_{ij}(\mathbf{a},t)\, \dot{\eta}_{,j}(\mathbf{a},t)] + r(\mathbf{a},t) = 0 \tag{10}$$

Equation (10) is exactly the equation governing the behavior of a thermoelastic body for $t \le t_0$, as given by (5).

Following the terminology of the calculus of variations, we conclude that $\mathfrak{L}_{t_0}$ *is a Lagrangian for thermoelasticity* (for $t \le t_0$). Note that in contrast with the method of adjoint fields, the Euler's equations of motion consist only of the thermoelastic equations of motion: there are no additional equations. This remarkable feature is entirely due to the symmetry of the convolution in time since if $E(t)$ represents the Euler-Lagrange expression at time t, and $\delta B(t)$ the variation of the appropriate fields at t, we have:

$$\int_0^{t_0} \{A(t)\, \delta B(t_0 - t) + A(t_0 - t)\, \delta B(t)\}\, dt = 2\int_0^{t_0} A(t)\, \delta B(t_0 - t)\, dt \tag{11}$$

Another consequence of (11) is that even though t_0 appears in a crucial way in (8), it does not enter in any way the equations (10), which was to be expected in as much as Euler's equations should be a statement about the time evolution of a given system, a statement that cannot possibly involve the future of this system.

It is actually possible to produce a set of "natural" boundary and initial conditions from the variation of A. Demanding (u_i,η) to vanish at time $t = 0$ and A to be stationary over all elements of $C^{\infty}(\bar{B} \times [0,t_0], \mathbb{R}^4)$ that vanish at $t = 0$, we obtain as boundary and initial conditions:

Boundary Conditions

$$\sigma_{ij}(\mathbf{a},t)\, n_i = 0, \quad \text{(traction free boundary)}$$

$$q_i(\mathbf{a},t)\, n_i = 0, \quad \text{(flux free boundary)} \tag{12}$$

n_i being the ith component of the outwardly directed normal to ∂B, the boundary of the body.

Initial Conditions

$$x_i(\mathbf{a},0) = a_i, \quad \text{(no initial displacement)}$$

$$\dot{x}_i(\mathbf{a},0) = 0, \quad \text{(no initial velocity)} \tag{13}$$

$$\eta(\mathbf{a},0) = 0$$

$$s(\mathbf{a},0) = 0, \quad \text{(no initial entropy)} \tag{14}$$

Note that the first equation of (14) is satisfied by choosing

$$\eta(\mathbf{a},t) = \int_0^t \tau(\mathbf{a},t')\, dt' \tag{15}$$

which determines η completely.

The system of equations (10), (13)-(15) is a well-posed, boundary-initial-value problem in x_i and η, provided that some regularity conditions are met by f_i and r, the external "loadings" (refer to Dafermos [15] for example). Once more the convolution is essential to obtain a meaningful set of boundary and initial conditions. This has to be contrasted with the classical Lagrangian for elasticity given by (9) which does not give rise to a well-posed initial boundary-value problem.

In the next section we generate material momenta from our Lagrangian density and we briefly examine the connection between these material momenta and the existence of path-independent integrals in thermoelasticity.

5 Conservation of the Material Momentum

The conservation laws can be derived in various ways. A familiar one involving use of the Lagrangian for the system under consideration consists of allowing a larger class of variations (with varying boundaries) and to impose the invariance conditions on the corresponding action integral (refer e.g., to Logan [16]). The alternative simpler procedure was suggested for an elastic continuum by Golebiewska Herrmann [7]. It involves the Lagrangian density only and consists of differentiating it with respect to independent

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

variables. The terms are grouped in such a way that the use of the equations of motion can be made. We follow that approach. In view of (8), this is a straightforward, though lengthy computation. After numerous integrations by parts, we obtain:

$$2\frac{d\mathcal{L}_{t_0}}{da_m} = 2\frac{\partial \mathcal{L}_{t_0}}{\partial a_m}\Big|_{\exp} - [(A),x_{i,m}]_{t_0}^{+} + [(B),\eta_{,m}]_{t_0}^{+}$$
$$- \partial_j\{[\sigma_{ij},x_{i,m}]_{t_0}^{+} - [q_j,\eta_{,m}]_{t_0}^{+}\} + \partial_t\{\rho(\mathbf{a})[\dot{x}_i,x_{i,m}]_{t_0}^{-}$$
$$- [s,\eta_m]_{t_0}^{-} - \tfrac{1}{2}\lambda_{ij}(\mathbf{a})[\eta_{,j},\eta_{,im}]_{t_0}^{-}\} \tag{16}$$

where (A) and (B) denote the left-hand sides of equations $(5a)$ and $(5b)$ and

$$\frac{\partial \mathcal{L}}{\partial_{a_m}}\Big|_{\exp}$$

refers to explicit differentiation of $\mathcal{L}_{t_0}$ with respect to a_i.

We define $B_{mj}(\mathbf{a},t,t_0)$ and $B_m(\mathbf{a},t,t_0)$ to be

$$B_{mj}(\mathbf{a},t,t_0) = -\tfrac{1}{2}[\sigma_{ij},x_{i,m}]_{t_0}^{+} + \tfrac{1}{2}[q_j,\eta_{,m}]_{t_0}^{-} - \mathcal{L}_{t_0}\delta_{jm}$$
$$B_m(\mathbf{a},t,t_0) = +\tfrac{1}{2}\rho(\mathbf{a})[\dot{x}_i,x_{i,m}]_{t_0}^{-} - \tfrac{1}{2}[s,\eta_{,m}]_{t_0}^{-}$$
$$- \tfrac{1}{4}\lambda_{ij}(\mathbf{a})[\eta_{,j},\eta_{,im}]_{t_0}^{-} \tag{17}$$

Then taking $(5a)$ and $(5b)$ into account, (16) reads as

$$\frac{\partial B_{mj}}{\partial a_j} + \frac{\partial B_m}{\partial t} = -\frac{\partial \mathcal{L}_{t_0}}{\partial a_m}\Big|_{\exp} \tag{18}$$

If we assume *homogeneity of the body and space uniformity of the mechanical and thermal loadings*, the explicit derivatives vanish and we are left with space-time, divergence-free quantities composed of B_{mj} and B_m. Following the terminology of Golebiewska Herrmann [7], we will refer to (18) with vanishing right-hand side as to a *material conservation law*. Introducing the symbolic notation $\mathfrak{B} = (B_{mj},B_m)$, the latter can be written shortly as

$$div_4\,\mathfrak{B} = 0 \tag{19}$$

The relation (18), however, has an even more general meaning, since

$$\frac{\partial \mathcal{L}_{t_0}}{\partial a_m}\Big|_{\exp}$$

will be different from zero in any part of the body where a defect or an inhomogeneity is present. In this respect (18) can be viewed as a balance of material momentum.

An essential difference with the elastic case is evidenced in (17). The material momentum $\mathfrak{B}$ is nonlocal in time. This is a direct consequence of the dissipation and it has profound implications on the possibility of generating path-independent integrals. Such integrals are usually obtained by application of the divergence theorem to a material balance law. A direct inspection of (19) shows that if we consider a domain of the form $Gx[0,t_0]$, where G is a subdomain of B, a significant simplification occurs due to the symmetry property of the convolution. Defining $I_{mj}(\mathbf{a},t_0)$ to be

$$\int_0^{t_0} B_{mj}(\mathbf{a},t,t_0)\,dt = \int_0^{t_0}\{\mathcal{L}_{t_0}\delta_{jm} - \sigma_{ij}(\mathbf{a},t)\,x_{i,m}(\mathbf{a},t_0-t)$$
$$+ q_j(\mathbf{a},t)\,\eta_{,m}(\mathbf{a},t_0-t)\}\,dt \tag{20}$$

we obtain from (19)

$$\int_{\partial G} I_{mj}(\mathbf{a},t_0)\,n_j d\sigma = -\Big\{\int_G (\rho\dot{x}_i(\mathbf{a},t)\,x_{i,m}(\mathbf{a},t_0-t)$$
$$-s(\mathbf{a},t)\,\eta_{,m}(\mathbf{a},t_0-t)$$
$$- \tfrac{1}{2}\lambda_{ij}\eta_{,j}(\mathbf{a},t)\,\eta_{,im}(\mathbf{a},t_0-t)\,d^3a\Big\}\Big|_{t=0}^{t=t_0} \tag{21}$$

where ∂G denotes the boundary of G, and n_j is the jth component of the outer normal to ∂G. We assume that this boundary is smooth enough for application of the divergence theorem.

Making use of relation (15), equation (21) becomes

$$\int_{\partial G} I_{mj}(\mathbf{a},t_0)\,n_j d\sigma = -\Big\{\int_G \rho\dot{x}_i(\mathbf{a},t)\,x_{i,m}(\mathbf{a},t_0 - t)\,d^3a\Big\}\Big|_{t=0}^{t=t_0}$$
$$-\int_G s(\mathbf{a},0)\,\eta_{,m}(\mathbf{a},t_0)\,d^3a \tag{22}$$

This last equation is the closest we can get to path independence: observe that the right-hand side depends only on $t = 0$ and $t = t_0$; it is the material analog of an impulse-momentum type relation.

In contrast with the elastic case, there is no true path independence even in a quasi-static case unless the initial entropy of the system happens to be equal to zero.

But the most striking difference with the classical path-independent integrals of fracture mechanics lies in the nonlocal character of all the quantities defined. It is only after specifying over which time length the system is observed that it becomes possible to devise for that time interval quantities such as $I_{mj}(\mathbf{a},t_0)$.

In a later paper we will show that, in analogy with the elastic case, the quantity

$$\int_S I_{mj}(\mathbf{a},t_0)\,n_j d\sigma$$

where S is the surface around a crack in a given body, and B is related to the total energy change of the system due to an infinitesimal translation in the mth direction of this crack, over the interval $[0,t_0]$. What this suggests is that it is not possible, as it is in elasticity, to determine the instantaneous force acting on the crack and to decide whether the crack will start propagating at a given instant in time, but that we should be able to determine if a crack will start propagating during any given time interval.

6 Concluding Remarks

The proposed Lagrangian differs from usual Lagrangians because of its nonlocality in time. In many physical theories there exists a strong relation between energy and time: in classical mechanics or field theory, invariance with respect to translation in time leads to conservation of energy; in quantum mechanics, Heisenberg's principle couples time and energy. Because the basic feature distinguishing a thermoelastic system for an elastic one lies in its dissipative character, it seems justified to adopt an expression where the time dependence differs from the one most often used in elasticity.

Let us emphasize once more that only one set of equations is derived (from the variational principle) in contradistinction to the formulation using adjoint fields.

In a following paper we will concentrate on quasi-static behavior and on the possible relation of material momentum to energy release rates.

7 Acknowledgment

The authors are grateful to Professors George Herrmann and Pawel Rafalski for most helpful discussions. This research was supported by Air Force Grant AF80-0149 to Stanford University.

References

1 Noether, E., "Invariant Variational Problems," *Trans. Th. Stat. Phys.*

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(trans. M. Tavel), Vol. 1, 1971, pp. 186–207. Also see *Nachr. Ges Gottingen* (Math-Phys. Klasse), Vol. 3, 1918, p. 235.

2 Eshelby, J. D., "The Force on an Elastic Singularity," *Phil. Trans. Roy. Soc.*, London, 1951, p. 87.

3 Günther, W., "Uber Einige Randintegrale der Elastomechanik," *Adhandlungen der Braunschweigischen Wissenschaftlichen Gesellschaft*, Vol. XIV, Verlag Friedr. Vieweg., Braunschweig, 1962.

4 Knowles, J. K., and Sternberg, E., "On a Class of Conservation Laws in Linearized and Finite Elastostatics," *Arch. Rat. Mech. Anal.*, Vol. 44, 1972, pp. 187–211.

5 Rice, J. R., "A Path-Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, 1968.

6 Budianksky, B., and Rice, J. R., "Conservation Laws and Energy Release Rates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, 1973, p. 201.

7 Golebiewska Herrmann, A., "On Conservation Laws of Continuum Mechanics," *Int. J. Solids and Structures*, Vol. 17, 1981, pp. 1–9.

8 Morse, Ph. M., and Feshbach, H., *Methods of Theoretical Physics*, McGraw-Hill, New York, 1953.

9 Tasi, J., and Herrmann, G., "Thermoelastic Dissipation in High-Frequency Vibrations of Crystal Plates," *J. Acoustical Soc. of Amer.*, Vol. 36, 1964, pp. 100–110.

10 Biot, M. A., *Variational Principles in Heat Transfer—A Unified Lagrangian Analysis of Dissipative Phenomena*, Clarendon Press, Oxford, 1970.

11 Rafalski, P., private communication.

12 Gurtin, M. E., "On a Path-Independent Integral for Thermoelasticity," *Int. J. Fracture*, Vol. 15, 1979, pp. 169–170.

13 Herrmann, G., "Some Applications of Invariant Variational Principles in Mechanics of Solids," *Variational Methods in the Mechanics of Solids*, Nemat-Nasser, S., ed., Pergamon Press, 1980, pp. 145–150.

14 Gurtin, M. E., "On a Path-Independent Integral for Elastodynamics," *Int. J. Fracture*, Vol. 12, 1976, pp. 643–644.

15 Dafermos, C., "On the Existence and Asymptotic Stability of Solution of the Equations of Thermoelasticity," *Arch. Rat. Mech. and Anal.*, Vol. 29, 1968, pp. 241–271.

16 Logan, J. D., *Invariant Variational Principles*, Academic Press, New York, 1977.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Stefan Problem With an Imperfect Thermal Contact at the Interface

L. N. Tao
Professor of Mechanics,
Department of Mechanics,
Mechanical, and Aerospace
Engineering,
Illinois Institute of Technology,
Chicago, Ill. 60616
Mem. ASME

The Stefan problem in a semi-infinite region with arbitrarily prescribed initial and boundary conditions, subject to a condition of the mixed type at the interface is investigated. To establish the exact solution of the problem, some new basic solutions of the heat equation are offered. Their mathematical properties are also supplied. The exact solutions of the temperatures in both phases and of the interfacial boundary are derived in infinite series. The existence and uniqueness of these series are considered and proved. It is also shown that these series are absolutely and uniformly convergent. Some concluding remarks about the differences between the present problem and the classical Stefan problem are given. Also the effect of a density discontinuity at the interface is discussed.

1 Introduction

The Stefan problem is concerned with the heat equation with a moving boundary which is not specified a priori. The first exact solution of the problem was established by Neumann (see [1]) in the 1860s and by Stefan [2] in 1889. They considered the problem of melting or freezing of a semi-infinite material with a constant initial temperature subject to another constant temperature at the surface. Solutions of the temperatures are expressed in error integrals of a similarity variable $x/t^{1/2}$ and the position of the interfacial boundary is proportional to $t^{1/2}$ where x is the space coordinate and t the time variable. Since then, there have been numerous studies intended to find exact solutions of the Stefan problems with more comprehensive initial and boundary conditions. Specific references may be found in many books [3–6] and in papers of two recent proceedings on moving boundary problems [7, 8]. There are also studies on the existence and uniqueness of the Stefan problems; references may be found in [3–8].

Only recently, the analytically exact solution of the Stefan problem in a semi-infinite region with arbitrarily prescribed initial and boundary conditions has been found [9, 10]. Solutions of the temperatures are expressed in infinite series of functions and polynomials in the error integral family (or more generally, of parabolic cylinder functions) and time t, and the interfacial boundary is a power series of $t^{1/2}$ (or of fractional powers of t). The exact solutions for the problem with arbitrary heat flux or with radiation (convective type) boundary condition have also been derived [11, 12].

In the classical Stefan problem, the conditions imposed at the interface are that the temperatures of both phases are at the freezing point and that the net heat flux across the interface is equal to the generation or absorption of latent heat of the phase change. Thus the temperatures at the interface are continuous while the temperature gradients are discontinuous. The requirement of continuity of temperatures at the interface might not be a valid one in some problems. There might be an imperfect thermal contact; an interface resistance is then formed such that the temperatures are not necessarily continuous across the interfacial boundary. For example, solidification of a binary alloy does not occur at the equilibrium temperature specified by the liquidus line of the phase diagram [13, 14]. Also, if the present problem is interpreted as a mass transfer problem of crystallization [15], there are many physical phenomena, such as nucleation, adsorption, surface diffusion, kink incorporation, and so on, by which a monomolecular layer at the interface is formed [16], causing the surface concentration to deviate from its saturation value.

In this paper we assume that there exists an imperfect thermal contact to form an interface resistance such that the temperatures of two phases are not required to be continuous across the phase boundary. The interface resistance is expressed in the form of a boundary condition of the third type, mimicking Newton's law of cooling. In other words, for solidification problems the heat flux from the liquid phase across the phase boundary is proportional to the temperature jump at the interface. The temperature of the liquid adjacent to the interfacial boundary is thus not known a priori. Moreover, when a temperature jump occurs, it results in an added term of energy exchange in the heat-balance equation.

The Stefan problems with the interface conditions of the third type have been studied, though not widely. Hill and Kotlow [17, 18] have investigated the existence and uniqueness of the problem and Ghez [15] has obtained some

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January, 1982; final revision, April, 1982. Paper No. 82-WA/APM-17.
Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

approximate solutions for the one-phase crystallization problem. However, no exact solution of the Stefan problem with these types of interface conditions has yet been found.

It is the purpose of this paper to establish the exact solution of the problem in a semi-infinite region with arbitrarily prescribed initial and boundary conditions. In the next section, we offer some new basic solutions to the heat equation which are needed in this study. Some mathematical properties of these functions are supplied. These functions were recently introduced by Tao [19] for some heat conduction problems in a fixed domain with a nonlinear boundary condition at its surface. In Section 3, the exact solutions of temperatures in both phases and of the interfacial boundary are established. They are expressed in infinite series; the existence and uniqueness of these series are considered and proved in Section 4. It is also shown there that the series are absolutely and uniformly convergent. We conclude the paper with some remarks pertaining to the problem. Some of the differences between the present problem and the classical Stefan problem are given. Also, the effect of a convective motion in the liquid phase caused by the density difference at the interface during phase transition is discussed.

2 Mathematical Preliminaries

Solutions of the heat equation

$$\alpha T_{xx} = T_t \tag{2.1}$$

may be expressed in terms of the repeated error integrals and time t, i.e.,

$$t^{n/2} i^n \operatorname{erfc}(\pm\xi), \quad \xi = x/\sqrt{4\alpha t} \tag{2.2}$$

Following [9, 11, 12], we introduce

$$\begin{aligned} E_n(\xi) &= \tfrac{1}{2}[i^n\operatorname{erfc}(-\xi) + i^n\operatorname{erfc}\xi] \\ F_n(\xi) &= \tfrac{1}{2}[i^n\operatorname{erfc}(-\xi) - i^n\operatorname{erfc}\xi] \\ G_n(\xi) &= \tfrac{1}{2}[i^n\operatorname{erfc}(-\xi) + (-1)^n i^n\operatorname{erfc}\xi] \\ &= \sum_0^{[n/2]} \xi^{n-2m}/[2^{2m} m!\,(n-2m)!] \end{aligned} \tag{2.3}$$

It is seen that $G_n(\xi)$ is a polynomial of degree n and the set is closely related to Widder's heat polynomials [20]. Mathematical properties of these functions can be derived from those of the repeated error integrals. For example, the integral representations are

$$\begin{aligned} \begin{bmatrix} E_n(\xi) \\ F_n(\xi) \end{bmatrix} &= \int_0^\infty [K(x-y,t;\alpha) \pm K(x+y,t;\alpha)] \\ &\quad \times (y/\sqrt{4\alpha t})^n dy/n! \\ G_n(\xi) &= \int_{-\infty}^\infty K(x-y,t;\alpha)(y/\sqrt{4\alpha t})^n dy/n! \end{aligned} \tag{2.4}$$

where

$$K(x,t;\alpha) = \exp(-x^2/4\alpha t)/\sqrt{4\pi\alpha t}$$

Also, we remark that at $t=0$ all three functions are unbounded, but their products with $(4\alpha t)^{n/2}$ have a definite limit,

$$\begin{aligned} (4\alpha t)^{n/2} E_n(\xi)|_{t=0} &= (4\alpha t)^{n/2} F_n(\xi)|_{t=0} \\ &= (4\alpha t)^{n/2} G_n(\xi)|_{t=0} = x^n/n! \end{aligned} \tag{2.5}$$

Because of the posed conditions at the interface, it is not possible to express the solutions to the problem in terms of only the foregoing functions; new basic solutions need to be generated. We offer the following lemma [19]:

If $u(x,t)$ is a solution of the one-dimensional heat equation (2.1), then

$$v(x,t) = e^{hx+h^2\alpha t} u(z,t)$$

is also a solution of the heat equation where

$$z = x + 2h\alpha t \quad \textit{and} \quad h = \textit{constant}.$$

Proof. We differentiate $v(x,t)$,

$$\begin{aligned} v_{xx} &= e^{hx+h^2\alpha t}(h^2 u + 2hu_z + u_{zz}) \\ v_t &= e^{hx+h^2\alpha t}(h^2\alpha u + 2h\alpha u_z + u_t) \end{aligned}$$

Since $\alpha u_{zz} = u_t$, we obtain $\alpha v_{xx} = v_t$. This completes the proof.

Therefore

$$\tau^n e^{\tau(2\xi+\tau)} i^n \operatorname{erfc}(\xi+\tau) \tag{2.6}$$

where $\tau = h\sqrt{\alpha t}, \quad \xi = x/\sqrt{4\alpha t}$

are also solutions of the heat equation. The x-derivative of

$$e^{\tau(2\xi+\tau)} i^n \operatorname{erfc}(\xi+\tau) \tag{2.7}$$

is

$$\begin{aligned} &\frac{h}{2\tau}\frac{\partial}{\partial\xi}[e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)] \\ &\quad = he^{\tau(2\xi+\tau)}\left[i^n\operatorname{erfc}(\xi+\tau) - \frac{1}{2\tau} i^{n-1}\operatorname{erfc}(\xi+\tau)\right] \end{aligned} \tag{2.8}$$

which is unbounded as $\tau\to 0$. This difficulty can be circumvented if we introduce

$$e^{\tau(2\xi+\tau)} i^n \operatorname{erfc}(\xi+\tau) - i^n\operatorname{erfc}\xi \tag{2.9}$$

Its x-derivative is

$$\begin{aligned} he^{\tau(2\xi+\tau)}\{&i^n\operatorname{erfc}(\xi+\tau) - \frac{1}{2\tau}[i^{n-1}\operatorname{erfc}(\xi+\tau) \\ &- e^{-\tau(2\xi+\tau)} i^{n-1}\operatorname{erfc}\xi]\} \end{aligned} \tag{2.10}$$

which is bounded and has a definite limit as $\tau\to 0$, since

$$\lim_{\tau\to 0} \frac{i^{n-1}\operatorname{erfc}(\xi+\tau) - e^{-\tau(2\xi+\tau)} i^{n-1}\operatorname{erfc}\xi}{2\tau} = -ni^n\operatorname{erfc}\xi$$

For subsequent uses, some mathematical properties of (2.7) are given in the following:

(*i*) Derivatives:

$$\begin{aligned} &\frac{\partial}{\partial\tau}[e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)] \\ &= e^{\tau(2\xi+\tau)}[2(\xi+\tau) i^n\operatorname{erfc}(\xi+\tau) - i^{n-1}\operatorname{erfc}(\xi+\tau)] \\ &\quad = -2(n+1)e^{\tau(2\xi+\tau)} i^{n+1}\operatorname{erfc}(\xi+\tau) \end{aligned} \tag{2.11}$$

$$\begin{aligned} &\frac{\partial^N}{\partial\tau^N}[e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)] \\ &\quad = (-2)^N[(N+n)!/n!]e^{\tau(2\xi+\tau)} i^{N+n}\operatorname{erfc}(\xi+\tau) \end{aligned}$$

(*ii*) Integral Representation:

$$e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau) = (2/\sqrt{\pi})\int_0^\infty z^n \exp[-(\xi+z)^2 - 2\tau z]dz/n! \tag{2.12}$$

This is a direct consequence of the lemma applied to the integral representation of $i^n\operatorname{erfc} z$.

(*iii*) Asymptotic Expansion:

$$e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau) = \pi^{-1/2} e^{-\xi^2}(\xi+\tau)^{n-1}\{1 + 0[(\xi+\tau)^{-2}]\} \tag{2.13}$$

This is obtained from the asymptotic expansion of $i^n\operatorname{erfc} z$.

(*iv*) Monotonicity:

For positive τ, $e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)$ is a strictly decreasing function bounded between $i^n\operatorname{erfc}\xi$ and 0. This assertion follows from

$$\begin{aligned} e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)|_{\tau=0} &= i^n\operatorname{erfc}\xi \\ e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)|_{\tau\to\infty} &= 0 \end{aligned} \tag{2.14}$$

and

$$\begin{aligned} &\frac{\partial}{\partial\tau}[e^{\tau(2\xi+\tau)} i^n\operatorname{erfc}(\xi+\tau)] \\ &\quad = -2(n+1)e^{\tau(2\xi+\tau)} i^{n+1}\operatorname{erfc}(\xi+\tau) < 0 \end{aligned}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(v) Inequality:

$$i^n\text{erfc}\xi \geq e^{\tau(2\xi+\tau)} i^n\text{erfc}(\xi+\tau) \geq 0 \tag{2.15}$$

for $\tau > 0$. It follows from

$$i^n\text{erfc}\xi - e^{\tau(2\xi+\tau)} i^n\text{erfc}(\xi+\tau) = (2/\sqrt{\pi})\int_0^\infty z^n e^{-(\xi+z)^2}(1-e^{-2\tau z})dz/n! \geq 0$$

3 Exact Solutions

To describe the problem, let us be specific and consider the solidification problem. The material is originally in a liquid state, occupying a half space $x > 0$, with an initial temperature $V(x)$ above the freezing temperature T_f. At its surface $x=0$, a temperature $U(t)$ below the freezing point is prescribed. Designate the solid and liquid phases of the material by subscripts I and II, respectively; then the temperatures satisfy

(i) Differential Equations:

$$\alpha_I(\partial^2 T_I/\partial x^2) = \partial T_I/\partial t \quad 0 < x < s(t)$$
$$\alpha_{II}(\partial^2 T_{II}/\partial x^2) = \partial T_{II}/\partial t \quad s(t) < x < \infty \tag{3.1}$$

where $s(t)$ is the position of the interfacial boundary;

(ii) Boundary and Initial Conditions:

$$T_I(0,t) = U(t), \quad T_{II}(x,0) = V(x) \tag{3.2}$$

(iii) Interface Conditions:

$$T_I(s,t) = T_f$$
$$(\partial T_{II}/\partial x)_s = h[T_{II}(s,t) - T_f] \tag{3.3}$$
$$k_I(\partial T_I/\partial x)_s - k_{II}(\partial T_{II}/\partial x)_s = \rho[L + c(T_{II} - T_f)_s](ds/dt)$$
$$s(0) = 0$$

where h is the heat transfer coefficient at the interface, ρ the density, c the specific heat of the liquid, and L the latent heat.

Two of these interface conditions are different from those of the classical Stefan problem, where temperatures are required to be continuous across the interface, i.e.,

$$T_{II}(s,t) = T_f$$
$$k_I(\partial T_I/\partial x)_s - k_{II}(\partial T_{II}/\partial x)_s = \rho L(ds/dt)$$

With an imperfect thermal contact between phases, temperatures of the present problem are not continuous across the interface. The liquid temperature is allowed to be above the freezing point and the heat flux out of the liquid phase is taken to be proportional to the temperature difference. Also, due to this temperature difference a heat balance of the phase transition results in an additional heat exchange at the interface. These more general conditions at the interface have been previously used by Hill and Kotlow [17, 18] in their studies on the existence and uniqueness of the problem.

To complete the description, the boundary and initial conditions need to be specified. We assume that they are arbitrary but expressible by power series,

$$U(t) = \sum_0 u_n (h^2\alpha_I t)^n/n!, \quad u_n = (d^n U/dt^n)_0/(h^2\alpha_I)^n \tag{3.4}$$
$$V(x) = \sum_0 v_n (hx)^n/n!, \quad v_n = (d^n V/dx^n)_0/h^n$$

With these two series in mind, we introduce the temperature solutions

$$T_I(x,t) = \sum_0 u_n (2\tau/\mu)^{2n} G_{2n}(\mu\xi) + \sum_0 a_n (2\tau/\mu)^n F_n(\mu\xi)$$
$$T_{II}(x,t) = \sum_0 v_n (2\tau)^n G_n(\xi) \tag{3.5}$$
$$+ \sum_0 b_n (2\tau)^n [i^n\text{erfc}\xi - e^{\tau(2\xi+\tau)} i^n\text{erfc}(\xi+\tau)]$$

where

$$\mu = \sqrt{\alpha_{II}/\alpha_I}, \quad \tau = h\sqrt{\alpha_{II} t}, \quad \xi = x/\sqrt{4\alpha_{II} t}$$

Clearly every term in these two temperature expressions is a solution of its respective heat equation. Furthermore, the prescribed boundary and initial conditions are satisfied:

$$T_I(0,t) = \sum_0 u_n (2\tau/\mu)^{2n} G_{2n}(0) = \sum_0 u_n (h^2\alpha_I t)^n/n! = U(t)$$
$$T_{II}(x,0) = \sum_0 v_n (2\tau)^n G_n(\xi)|_{t=0} = \sum_0 v_n (hx)^n/n! = V(x)$$

To complete the solutions we need to find the sets of coefficients $\{a_n\}$ and $\{b_n\}$ and the interfacial position $s(t)$ to satisfy the interface conditions (3.3). As in many moving boundary problems, the interfacial position $s(t)$ must be expressed as a power series of $t^{1/2}$, not t, in order to satisfy these conditions. Accordingly, we introduce

$$s(t) = 2\sqrt{\alpha_{II} t}\sum_0 \lambda_n \tau^n = (2\tau/h)\sum_0 \lambda_n \tau^n \tag{3.6}$$

Also we write

$$\eta(t) = s(t)/(4\alpha_{II} t)^{1/2} = \sum_0 \lambda_n \tau^n$$
$$\zeta(t) = [s(t) + 2h\alpha_{II} t]/(4\alpha_{II} t)^{1/2} = \sum_0 \lambda_n^* \tau^n \tag{3.7}$$

where

$$\lambda_1^* = \lambda_1 + 1, \quad \lambda_n^* = \lambda_n \ (n \neq 1)$$

Substitutions of (3.5), (3.6), and (3.7) into (3.3) and evaluations at $\tau = 0$ yield

$$a_0 = (T_f - u_0)/\text{erf}(\mu\lambda_0), \quad b_0 = (T_f - v_0 + v_1)/\text{erfc}\lambda_0 \tag{3.8}$$

and λ_0 satisfies

$$\mu\lambda_0 e^{\mu^2\lambda_0^2}\text{erf}(\mu\lambda_0) = [k_I(T_f - u_0)/\pi^{1/2}\rho\alpha_{II}]/[L + c(v_0 - T_f)]$$

With these coefficients, it is seen from (3.5) that at the beginning of solidification the temperature of the liquid adjacent to the interface is $V(0)$, not T_f. Thus the liquid phase does not undergo a sudden change of temperature at the interface as in the case of the classical Stefan problem where a sudden change from $V(0)$ to T_f occurs [9, 10].

To determine the remaining coefficients a_n, b_n, and λ_n, we may expand all terms of the interface equations (3.3) in series of τ and match the coefficients of equal powers. The process is somewhat cumbersome; we prefer to differentiate them with respect to τ and evaluate them at $\tau = 0$. Also, to avoid negative powers of τ, we differentiate the heat-balance equation after it is multiplied by τ. Differentiations yield

$$D_\tau^N[T_I(s,t) - T_f]_0 = 0$$
$$D_\tau^N[(\partial T_{II}/\partial x)_s - h(T_{II} - T_f)_s]_0 = 0 \tag{3.9}$$
$$k_I D_\tau^N[\tau(\partial T_I/\partial x)_s]_0 - k_{II} h D_\tau^N\{\tau[T_{II}(s,t) - T_f]\}_0$$
$$= \rho h\alpha_{II} D_\tau^N\{[L + c(T_{II} - T_f)_s]\sum_0 (n+1)\lambda_n \tau^n\}_0$$

where

$$(\partial T_I/\partial x)_s = h\sum_1 u_n (2\tau/\mu)^{2n-1} G_{2n-1}(\mu\eta)$$
$$+ h\sum_0 a_n (2\tau/\mu)^{n-1} E_{n-1}(\mu\eta)$$

$$[\partial T_{II}/\partial x - h(T_{II} - T_f)]_s = hT_f + h\sum_0 (v_{n+1} - v_n)(2\tau)^n G_n(\eta)$$
$$- h\sum_0 b_n (2\tau)^n i^n\text{erfc}\eta$$
$$+ h\sum_1 b_n (2\tau)^{n-1}[e^{\zeta^2 - \eta^2} i^{n-1}\text{erfc}\zeta - i^{n-1}\text{erfc}\eta] \tag{3.10}$$

When $N = 1$, we find

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$a_1 = -\mu^2\Phi_1(\mu\lambda_0)a_0\lambda_1/2F_1(\mu\lambda_0),\ \Phi_1(z)=d(\text{erf } z)/dz$$

$$b_1=[\Phi_1(\lambda_0)b_0\lambda_1+2(v_2-v_1)\lambda_0]/4i\text{erfc }\lambda_0$$

$$(2/k_{II})\{\rho\alpha_{II}[L+c(v_0-T_f)]$$
$$+[\mu^2k_IE_2(\mu\lambda_0)\Phi_1(\mu\lambda_0)/F_1(\mu\lambda_0)]a_0\}\lambda_1$$
$$=T_f-v_0-c(v_1\lambda_0+2b_0i\text{erfc }\lambda_0)\lambda_0 \tag{3.11}$$

The remaining coefficients can be found with the aid of a formula for the material time derivative of arbitrary order [21]. In the one-dimensional case where $f=f(x,t)$ and $x=x(t)$, we have

$$D_t^Nf(x,t)=\sum_{n=0}^{N}\sum_{m=0}^{N-n}(N!/n!)Z_m^{N-n}(x)\partial_t^n\partial_x^mf(x,t) \tag{3.12}$$

and

$$Z_m^n(x)=\sum_{\alpha_r}\prod_{r=1}^{n}(D_t^rx/r!)^{\alpha_r}/\alpha_r!$$

The sum Σ_{α_r} is extended to all multinomial coefficients, i.e., all non-negative integers which satisfy

$$\sum_1^n\alpha_r=m,\quad \sum_1^n r\alpha_r=n$$

Multinomial coefficients α_r are known and have been tabulated [22].[1] Also, this formula, when f does not depend on t explicitly, reduces to Faa de Bruno's formula for the derivative of a function of a function [22, 23]

$$D_t^Nf(x)=N!\sum_{n=0}^{N}Z_n^N(x)\partial_x^nf(x) \tag{3.13}$$

Using (3.12), we obtain

$$D_\tau^N[(2\tau/\mu)^nG_n(\mu\eta)]_0=\sum_{r=0}^{N}\sum_{m=0}^{N-r}(N!/r!)(2/\mu)^nZ_m^{N-r}(\eta)$$
$$\times\partial_\tau^r\partial_\eta^m[\tau^nG_n(\mu\eta)]|_0=N!A_n^N(\mu)$$

$$D_\tau^N[(2\tau/\mu)^ni^n\text{erfc}(\mu\eta)]_0=N!B_n^N(\mu) \tag{3.14}$$

$$D_\tau^N[(2\tau/\mu)^nE_n(\mu\eta)]_0=N!\{A_n^N(\mu)+\tfrac{1}{2}[1-(-1)^n]B_n^N(\mu)\}$$

$$D_\tau^N[(2\tau/\mu)^nF_n(\mu\eta)]_0=N!\{A_n^N(\mu)-\tfrac{1}{2}[1+(-1)^n]B_n^N(\mu)\}$$

where

$$A_n^N(\mu)=2^n\sum_0^{n^*}\mu^{m-n}Z_m^{N-n}(\eta_0)G_{n-m}(\mu\lambda_0),\quad n^*=\min(n,N-n)$$

$$B_n^N(\mu)=2^n\sum_0^{N-n}(-1)^m\mu^{m-n}Z_m^{N-n}(\eta_0)i^{n-m}\text{erfc}(\mu\lambda_0) \tag{3.15}$$

$$Z_m^n(\eta_0)=\sum_{\alpha_r}\prod_1^n(\lambda_r)^{\alpha_r}/\alpha_r!$$

Here we have used

$$i^{-m}\text{erfc }z=(-1)^{m+1}\Phi_m(z)=(2/\sqrt{\pi})e^{-z^2}H_{m-1}(z)$$
$$(m>0) \tag{3.16}$$

where $\Phi_m(z)$ and $H_m(z)$ are the iterated derivative of erf z and the Hermite polynomial, respectively. Tabulated values of $\Phi_m(z)$ and $H_m(z)$ are known [22, 24].

Similarly, we find

$$D_\tau^N[(2\tau)^ne^{\zeta^2-\eta^2}i^n\text{erfc }\zeta]_0$$

$$=N!\sum_{m=0}^{N-n}2^nD_\tau^m(e^{-\eta^2})_0D_\tau^{N-n-m}(e^{\zeta^2}i^n\text{erfc }\zeta)_0 \tag{3.17}$$

$$=N!C_n^N$$

where

$$C_n^N=\sum_{m=0}^{N-n}\sum_{p=0}^{m}\sum_{r=0}^{N-n-m}(-1)^{p+r}2^{n+r}[(n+r)!/r!]$$

$$Z_p^m(\eta_0)Z_r^{N-n-m}(\zeta_0)H_p(\lambda_0)i^{n+r}\text{erfc}(\lambda_0)$$

$$Z_m^n(\zeta_0)=\sum_{\alpha_r}\prod_1^n(\lambda_0^*)^{\alpha_r}/\alpha_r! \tag{3.18}$$

Substituting these formulas in the interface equations (3.9), we obtain the recurrence formulas for $N\geq 2$

$$\sum_0^{[N/2]}u_nA_{2n}^N(\mu)+\sum_0^Na_nA_n^N(\mu)-\sum_0^{[N/2]}a_{2n}B_{2n}^N(\mu)=0$$

$$\sum_0^N(v_{n+1}-v_n)A_n^N(1)-\sum_0^Nb_nB_n^N(1)$$
$$+\sum_1^Nb_n[C_{n-1}^N-B_{n-1}^N(1)]=0$$

$$(k_I/k_{II})\Big[\sum_1^{[N/2]}u_nA_{2n-1}^{N-1}(\mu)+\sum_1^Na_nA_{n-1}^{N-1}(\mu)$$
$$+\sum_0^{[N/2]}a_{2n}B_{2n-1}^{N-1}(\mu)\Big]-\sum_1^{N-1}v_nA_n^{N-1}(1)+\sum_1^{N-1}b_n[C_n^{N-1}$$
$$-B_n^{N-1}(1)]=(N+1)(\rho\alpha_{II}/k_{II})\Big\{[L+c(v_0-T_f)]\lambda_N$$
$$+\sum_{r=0}^{N-1}(r+1)\lambda_r\Big[\sum_{n=0}^{N-r}v_nA_n^{N-r}(1)+\sum_{n=0}^{N-r}b_n(B_n^{N-r}(1)-C_n^{N-r})\Big]\Big\} \tag{3.19}$$

From this set of equations, we may determine, starting from $N=2$ sequentially, all coefficients of a_n, b_n, and λ_n. This completes formally the series solutions of the problem.

4 Existence, Uniqueness, and Convergence

In this section we want to show that all coefficients a_n, b_n, and λ_n of the series solutions exist and are unique and that the series solutions of the problem are absolutely and uniformly convergent.

The coefficients a_0, b_0, and λ_0 are determined by (3.8). For positive λ_0 (3.8) has only one root [25], since the *L.H.S.* is strictly increasing from 0 to ∞. Although a negative λ_0 exists, it is physically inadmissible, since it corresponds to the interface moving initially outside of the material region. Thus all three coefficients of the zeroth order exist and are uniquely determined. It is also seen from (3.11) that the coefficients of the first order exist and are uniquely determined, since the coefficient of λ_1 in the *L.H.S.* of the third equation is always finite and positive.

To discuss the existence and uniqueness of the remaining coefficients, we observe that the only constants $A_n^N(\mu)$, $B_n^N(\mu)$, and C_n^N which contain λ_N are

$$B_0^N(\mu)=-\mu\lambda_N\Phi_1(\mu\lambda_0)$$
$$B_1^{N-1}(\mu)=-\mu^2\lambda_N\Phi_2(\mu\lambda_0)/2$$
$$C_0^N=-\lambda_N\Phi_1(\lambda_0)$$

Substituting these constants and

$$A_N^N(\mu)=(2/\mu)^NG_N(\mu\lambda_0)$$
$$B_N^N(\mu)=(2/\mu)^Ni^N\text{erfc}(\mu\lambda_0)$$
$$C_{N-1}^N-B_{N-1}^N(1)=-2^Ni^N\text{erfc}\lambda_0$$

in the first two equations of (3.19), we obtain

$$a_N=-[\mu^{N+1}\Phi_1(\mu\lambda_0)/2^NF_N(\mu\lambda_0)]a_0\lambda_N+\cdots$$
$$b_N=-[\Phi_1(\lambda_0)/2^{N+1}i^N\text{erfc}\lambda_0]b_0\lambda_N+\cdots$$

where the unwritten parts are independent of λ_N. Using these coefficients and the identities

[1] All multinomial coefficients for $n\leq 10$ are listed in [22]. For larger n or direct computation, a calculator program has been prepared by the author and has been deposited in the Texas Instruments' Professional Program Exchange.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\Phi_2(z) = -2z\Phi_1(z)$$

$$2nE_n(z) = E_{n-2}(z) + 2zF_{n-1}(z)$$

we find from the third equation of (3.19)

$$(N+1)\left\{\rho\alpha_{II}[L+c(v_0-T_f)] + \frac{\mu^2 k_I \Phi_1(\mu\lambda_0)E_{N+1}(\mu\lambda_0)}{\operatorname{erf}(\mu\lambda_0)F_n(\mu\lambda_0)}(T_f-u_0)\right\}\lambda_N = \cdots$$

where the *R.H.S.* contains only u_N and v_N and other coefficients of lower orders. From the fact that $E_n(\lambda)$, $F_n(\lambda)$, erf λ, and $\Phi_1(\lambda)$ are all positive when $\lambda>0$ and the hypotheses $v_0>T_f>u_0$, we conclude that λ_N exists and is uniquely determined. It in turn shows that a_N and b_N are uniquely determined. Thus all three sets of coefficients $\{a_n\}$, $\{b_n\}$, and $\{\lambda_n\}$ exist and are uniquely determined.

To discuss the convergence of the series solutions, let us first examine the interface conditions. Using the integral representations (2.4), we have

$$\left|\sum_0 u_n(2\tau/\mu)^{2n}G_{2n}(\mu\xi)\right| = \left|\int_{-\infty}^{\infty}\sum_0 u_n(hy)^{2n}K(x-y,t;\alpha_I)\,dy/n!\right| = \left|\int_{-\infty}^{\infty}U(\sqrt{y/\alpha_I})K(x-y,t;\alpha_I)\,dy\right| \le \max|U(t)|$$

$$\left|\sum_0 (v_{n+1}-v_n)(2\tau)^n G_n(\eta)\right| = \left|\int_{-\infty}^{\infty}\sum_0 (v_{n+1}-v_n)(hy)^n K(x-y,t;\alpha_{II})\,dy/n!\right| = \left|\int_{-\infty}^{\infty}\frac{1}{h}\left[\frac{dV(y)}{dy}-hV(y)\right]K(x-y,t;\alpha_{II})\,dy\right| \le \max|dV(x)/dx - hV(x)|/h$$

Since the coefficients λ_n exist, we denote s^* to be the nonzero interfacial position at time $\tau^*>0$. Then using the foregoing inequalities in (3.5) and (3.10), we conclude from the first two equations of (3.3) that

$$\left|\sum a_n(2\tau^*/\mu)^n F_n(\mu\eta^*)\right|$$

$$\left|\sum b_n(2\tau^*)^n\left[i^n\operatorname{erfc}\eta^* - \frac{1}{2\tau^*}(i^{n-1}\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^{n-1}\operatorname{erfc}\zeta^*)\right]\right|$$

are bounded, where the asterisk denotes evaluation at the interface s^*. Hence there exist some constants M_1 and M_2, independent of n, such that

$$|a_n| \le M_1/[(2\tau^*/\mu)^n F_n(\mu\eta^*)]$$

$$|b_n| \le M_2/\{(2\tau^*)^n\left[i^n\operatorname{erfc}\eta^* + \frac{1}{2\tau^*}(i^{n-1}\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^{n-1}\operatorname{erfc}\zeta^*)\right]$$

$$\le M_2/[(2\tau^*)^n i^n\operatorname{erfc}\eta^*]$$

$$\le M_2/[(2\tau^*)^n(i^n\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^n\operatorname{erfc}\zeta^*)]$$

Also, we observe that for $\tau<\tau^*$ and $\xi<\eta^*$

$$F_n(\mu\eta^*) > F_n(\mu\xi) > 0$$

since $i^n\operatorname{erfc}\xi$ is monotonically decreasing and $i^n\operatorname{erfc}(-\xi)$ is monotonically increasing. For $\tau<\tau^*$ and $\xi>\eta^*$

$$e^{\tau(2\xi+\tau)}i^n\operatorname{erfc}(\xi+\tau) > e^{\tau^*(2\xi+\tau^*)}i^n\operatorname{erfc}(\xi+\tau^*) > 0$$

$$i^n\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^n\operatorname{erfc}\zeta^* > i^n\operatorname{erfc}\xi - e^{\tau^*(2\xi+\tau^*)}i^n\operatorname{erfc}(\xi+\tau^*) > 0$$

These two inequalities follow directly from the integral representation of the repeated error integral and (2.12),

$$e^{\tau(2\xi+\tau)}i^n\operatorname{erfc}(\xi+\tau) - e^{\tau^*(2\xi+\tau^*)}i^n\operatorname{erfc}(\xi+\tau^*) = (2/\sqrt{\pi})\int_0^{\infty} z^n e^{-(\xi+z)^2}(e^{-2\tau z}-e^{-2\tau^* z})\,dz/n! > 0$$

$$i^n\operatorname{erfc}\eta^* - i^n\operatorname{erfc}\xi - e^{\zeta^{*2}-\eta^{*2}}i^n\operatorname{erfc}\zeta^* + e^{\tau^*(2\xi+\tau^*)}i^n\operatorname{erfc}(\xi+\tau^*) = (2/\sqrt{\pi})\int_0^{\infty} z^n[e^{-(\eta^*+z)^2} - e^{-(\xi+z)^2}](1-e^{-2\tau^* z})\,dz/n! > 0$$

To establish the convergence of the temperature solutions, we need only to show that the second parts of (3.5) are convergent. For $\tau<\tau^*$ and $\xi<\eta^*$

$$\left|\sum a_n(2\tau/\mu)^n F_n(\mu\xi)\right| \le M_1\sum\frac{F_n(\mu\xi)}{F_n(\mu\eta^*)}\left(\frac{\tau}{\tau^*}\right)^n \le M_1\sum(\tau/\tau^*)^n$$

and for $\tau<\tau^*$ and $\xi>\eta^*$

$$\left|\sum b_n(2\tau)^n[i^n\operatorname{erfc}\xi - e^{\tau(2\xi+\tau)}i^n\operatorname{erfc}(\xi+\tau)]\right| \le M_2\sum\frac{i^n\operatorname{erfc}\xi - e^{\tau(2\xi+\tau)}i^n\operatorname{erfc}(\xi+\tau)}{i^n\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^n\operatorname{erfc}\zeta^*}\left(\frac{\tau}{\tau^*}\right)^n$$

$$\le M_2\sum\frac{i^n\operatorname{erfc}\xi - e^{\tau^*(2\xi+\tau^*)}i^n\operatorname{erfc}(\zeta+\tau^*)}{i^n\operatorname{erfc}\eta^* - e^{\zeta^{*2}-\eta^{*2}}i^n\operatorname{erfc}\zeta^*}\left(\frac{\tau}{\tau^*}\right)^n \le M_2\sum(\tau/\tau^*)^n$$

They are majorized by geometric series; the series are thus absolutely and uniformly convergent. This implies that both temperature solutions are absolutely and uniformly convergent in their respective domains.

The convergence of the interfacial position $s(t)$ can be studied from the heat-balance equation. Since both temperature terms are bounded, the heat-balance equation is an ordinary differential equation of the first order for $s(t)$. The existence, uniqueness, and convergence are well known [26]. Thus the series solution of the interfacial position $s(t)$ is also bounded for all finite t.

5 Concluding Remarks

In this paper we have established an exact solution of the Stefan problem with arbitrary initial and boundary conditions subject to a condition of the mixed type at the interface. There are some differences between the present problem and the classical Stefan problem. When the temperature of the liquid is initially above the freezing point, there is no sudden change of its temperature at the initiation of solidification, in contrast to the classical problem where a sudden change of temperature from $V(0)$ to T_f occurs.

Another difference is the velocity of the interfacial boundary. The initial velocity of the present problem is specified by the second equation of (3.8); the corresponding λ_0 of the classical Stefan problem [9, 10] in the notation of this paper is given by

$$\mu\lambda_0 e^{\mu^2\lambda_0^2}\operatorname{erf}(\mu\lambda_0) = \frac{k_I(T_f-u_0)}{\pi^{1/2}\rho\alpha_{II}}\left[L+c(v_0-T_f)\frac{e^{-\lambda_0^2}}{\pi^{1/2}\lambda_0\operatorname{erfc}\lambda_0}\right]^{-1}$$

Since for positive λ_0,

$$\pi^{1/2}\lambda_0 e^{\lambda_0^2}\operatorname{erfc}\lambda_0$$

is monotonically increasing from 0 to 1 [25], we readily

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

conclude that the initial velocity of the interfacial boundary under the same initial and boundary conditions is slower in the present problem.

In the classical Stefan problem with constant initial and boundary conditions, the temperature solutions are expressible by a similarity variable $x/t^{1/2}$ in finite terms and the position of the interfacial boundary is proportional to $t^{1/2}$—the well-known Neumann solution. When the boundary condition of constant temperature is replaced by a constant heat flux at the surface, there is no similarity solution. A similarity solution exists only when the heat flux is proportional to $t^{-1/2}$ [27]. In the present problem with a mixed condition at the interface, no similarity solution can be found. This is due to the fact that a boundary condition of the mixed type does not permit a similarity solution.

In this paper we have followed the usual assumption that the liquid and solid phases of the material have the same density. If this assumption is lifted, then there is a convective motion in the liquid caused by the density jump at the interface. In another paper [28], it is shown that the Stefan problem including the density jump at the interface can be transformed by some appropriate changes of the independent variables and material constants to an associated problem in which both liquid and solid phases behave as if they had the same density. Hence, solutions of the problems without a density difference can be utilized for those with a density jump at the interface. This transformation is also valid for the present problem. We refer to the original paper for details.

Acknowledgment

This material is based on work supported by the National Science Foundation under Grant CME 79-21350.

References

1 Riemann, G. F. B., and Weber, H., *Die Partiellen Differentialgleichungen der Mathematischen Physik*, Bd. 2, 5-fl. Aufl., Braunschweig, 1912.

2 Stefan, J., "Über Einige Probleme der Theorie der Wärmeleitung," *S.-B. Wien. Akad. Mat. Natur.*, Vol. 98, 1889, pp. 473-484.

3 Carslaw, H. S., and Jaeger, J. C., *Conduction of Heat in Solids*, 2nd ed., Clarendon Press, Oxford, 1959.

4 Crank, J., *The Mathematics of Diffusion*, 2nd ed., Clarendon Press, Oxford, 1975.

5 Rubinstein, L. I., "The Stefan Problem," *Transl. Math. Monog. No. 27, Am. Math. Soc.*, Providence, R. I., 1971.

6 Özisik, M. N., *Heat Conduction*, Wiley, New York, 1980.

7 Ockendon, J. R., and Hodgkins, W. R., Eds., *Moving Boundary Problems in Heat Flow and Diffusion*, Clarendon Press, Oxford, 1975.

8 Wilson, D. G., Solomon, A. D., and Boggs, P. T., Eds., *Moving Boundary Problems*, Academic Press, New York, 1978.

9 Tao, L. N., "The Stefan Problem With Arbitrary Initial and Boundary Conditions," *Quart. Appl. Math.*, Vol. 36, 1978, pp. 223-233.

10 Tao, L. N., "The Analyticity of Solutions of the Stefan Problem," *Arch. Rational Mech. Anal.*, Vol. 72, 1980, pp. 285-301.

11 Tao, L. N., "On Free Boundary Problems With Arbitrary Initial and Flux Conditions," *Z. Angew. Math. Phys.*, Vol. 30, 1979, pp. 416-426.

12 Tao, L. N., "Free Boundary Problems With Radiation Boundary Conditions," *Quart. Appl. Math.*, Vol. 37, 1979, pp. 1-10.

13 Chalmers, B., *Principles of Solidification*, Wiley, New York, 1964.

14 Tao, L. N., "On Solidification of a Binary Alloy," *Quart. J. Mech. Appl. Math.*, Vol. 33, 1980, pp. 211-225.

15 Ghez, R., "Expansions in Time for the Solution of One-Dimensional Stefan Problems of Crystal Growth," *Int. J. Heat Mass Transfer*, Vol. 23, 1980, pp. 425-432.

16 Landau, L. D., and Lifshitz, E. M., *Statistical Physics*, Academic Press, New York, 1958.

17 Hill, C. D., and Kotlow, D. B., "Classical Solutions in the Large of a Two-Phase Free Boundary Problem, I," *Arch. Rational Mech. Anal.*, Vol. 45, 1972, pp. 63-78.

18 Hill, C. D., and Kotlow, D. B., "Classical Solutions in the Large of a Two-Phase Free Boundary Problem, II," *Arch. Rational Mech. Anal.*, Vol. 47, 1972, pp. 369-379.

19 Tao, L. N., "Heat Conduction With Nonlinear Boundary Condition," *Z. Angew. Math. Phys.*, Vol. 32, 1981, pp. 144-155.

20 Widder, D. V., *The Heat Equation*, Academic Press, New York, 1975.

21 Tao, L. N., "On the Material Time Derivative of Arbitrary Order," *Quart. Appl. Math.*, Vol. 36, 1978, pp. 323-324.

22 Abramowitz, M., and Stegun, I. A., Eds., *Handbook of Mathematical Functions*, NBS AMS#55, U. S. Government Printing Office, Washington, D.C., 1964.

23 Jordan, C., *Calculus of Finite Difference*, Chelsea, New York, 1965.

24 Jahnke, E., and Emde, F., *Tables of Functions*, Dover, New York, 1945.

25 Tao, L. N., *Dynamics of Growth or Dissolution of a Gas Bubble*," *J. Chem. Phys.*, Vol. 69, 1978, pp. 4189-4194.

26 Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

27 Tao, L. N., "The Exact Solutions of Some Stefan Problems With Prescribed Heat Flux," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 732-736.

28 Tao, L. N., "On Solidification Problems Including the Density Jump at the Moving Boundary," *Quart. J. Mech. Math.*, Vol. 32, 1979, pp. 175-185.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Constitutive Model of Cyclic Plasticity With a Nonhardening Strain Region

N. Ohno[1]

School of Energy Engineering,
Toyohashi University of Technology,
Tempaku-cho, Toyohashi 440, Japan

By introducing the concept of a nonhardening region in the plastic strain space, a constitutive model is proposed for cyclic plastic loadings between variable, as well as fixed, strain limits. It is assumed that the isotropic hardening of materials does not occur when the plastic strain point moves inside this region after a load reversal. The region expands and translates as cyclic straining proceeds, and when strain limits are fixed, it eventually occupies the cyclic range of plastic strain so as to describe the saturation of cyclic hardening. The phenomena of cyclic relaxation and cyclic creep are also taken into account in the formulation. In the simple case of a linear hardening material, the present theory is verified by comparing predictions with experimental results on type 304 stainless steel under torsional cyclings between variable, as well as fixed, strain limits at room temperature.

1 Introduction

Cyclic plasticity of metals and alloys has been investigated experimentally by many researchers [1-7]. It has been observed that a stable hysteresis loop can be approached after a period of cyclic hardening or softening under cycling between fixed strain limits. This phenomenon is referred to as the stabilization of the hysteresis loop. However, if the strain amplitude is changed after stabilization of the hysteresis loop, additional cyclic hardening or softening is observed until the hysteresis loop is stabilized again [2, 4, 6]. Therefore, the stabilization of the hysteresis loop is not permanent. Other important phenomena in cyclic plasticity are the relaxation of mean stress under cycling between unequal strain limits (cyclic relaxation) and the accumulation of plastic strain under cycling between unequal stress limits (cyclic creep).

Isotropic hardening [8] and/or kinematic hardening [9] of materials are assumed in classical theories of plasticity. Then, the yield condition may be expressed as follows:

$$f = \phi(s_{ij} - R_{ij}) - \kappa^2(q) = 0 \tag{1}$$

where s_{ij} and R_{ij} are second-rank tensors indicating the deviatoric stress and the translation of the yield surface, respectively, and q is a scalar proportional to the accumulated plastic strain or the plastic work. If it is assumed that the yield surface has a saturation size (i.e., $\kappa(\infty) =$ finite), such theories may describe the stabilization phenomenon of a hysteresis loop. It is evident, however, that this assumption leads to the permanent stabilization of a hysteresis loop except for the contribution of kinematic hardening. Hence, the accumulated plastic strain or the plastic work is not an appropriate variable to describe the isotropic hardening in cyclic plasticity, especially when the cyclic range varies. Moreover, when strain limits are fixed at arbitrary values, this remark holds true, because the saturation level of stress amplitude becomes identical for all strain amplitudes except for the contribution of kinematic hardening.

The preceding discussion suggests that the isotropic hardening of materials in cyclic plasticity should be described in terms of a variable that takes a finite saturation value depending on the strain amplitude in the stabilized state of a

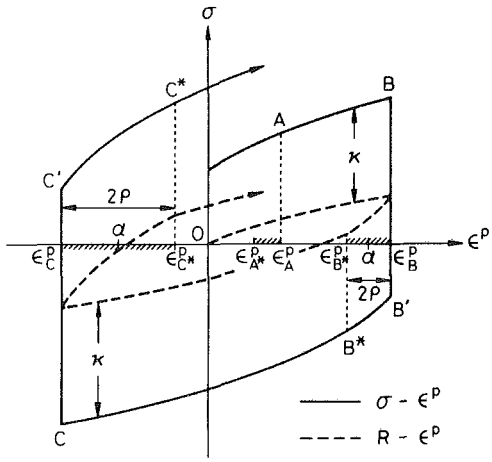


Fig. 1 Evolution of nonhardening strain region in uniaxial state of stress

[1] Currently, Visiting Scholar, Harvard University, Division of Applied Sciences, Cambridge, Mass. 02138.

Contributed by the Applied Mechanics Division for Publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper shoud be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October, 1981, final revision, May, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

hysteresis loop. Recently, by assuming the dependence of cyclic hardening on the maximum plastic strain range, Chaboche et al. [6] proposed such a variable to describe the cyclic behavior of type 316 stainless steel under variable strain amplitude. This assumption was employed by Dafalias [10] in the context of a bounding surface theory. However, since the original assumption was obtained on the basis of the experimental results at zero mean strain, further investigations will be necessary to verify it under more general conditions (see Section 3).

Other theories for the saturation of cyclic hardening or softening were presented by Bruhns [11], Eisenberg [12], and Bodner et al. [13]. The endochronic theory was also applied to this problem [14, 15]. Regarding the translation of the yield surface, on the other hand, Mróz et al. [16] generalized the rule of Prager [9] and that of Eisenberg and Phillips [17] to express the phenomena of cyclic relaxation and cyclic creep. An evanescent strain-memory effect was considered by Chaboche [18] to describe the nonlinear kinematic hardening. Translation rules using two stress surfaces were also proposed [19–22]. In these works, however, the aforementioned prescription of isotropic hardening was not discussed, and cyclic straining between variable strain limits was not considered either.

In the present study, a new constitutive model of cyclic plasticity is proposed by specifying a nonhardening region of plastic strain inside which the isotropic hardening does not occur. A possible interpretation of this region is given from the microstructural point of view. Then, the evolution equation of the nonhardening region and those of the isotropic and kinematic hardening variables are formulated. It is shown that this isotropic hardening variable takes a finite saturation value depending on the strain amplitude under fixed cyclic straining, and thus the present theory is applicable to cyclic straining between variable limits as well. In the case of a linear hardening material, the predictions are compared with experimental results on type 304 stainless steel at room temperature to verify the present theory.

2 Constitutive Assumption and Its Formulation

Nonhardening Strain Region. The gradual change of stress amplitude toward an asymptotic value under fixed cyclic straining suggests that the development of isotropic hardening decreases with increase in the number of cycles, resulting in only kinematic hardening in the stabilized state of a hysteresis loop. In order to describe such changes of internal state, let us assume a plastic strain region inside which the isotropic hardening of materials does not develop. Then, the cyclic straining within this region does not induce any further cyclic hardening, and hence it may be called the plastic region of cyclic nonhardening, or simply, the nonhardening strain region.

The region expands and translates as plastic deformation proceeds, as shown by the shaded portions in Fig. 1. When plastic deformation proceeds to the points A and B on the diagram of stress σ and plastic strain ϵ^p, it expands and translates to occupy the plastic strain ranges $[\epsilon^p{}_{A'}, \epsilon^p{}_A]$ and $[\epsilon^p{}_{B'}, \epsilon^p{}_B]$, respectively. After the load reversal at $\epsilon^p = \epsilon^p{}_B$, on the other hand, it remains unchanged as long as the plastic strain point ϵ^p is located inside the nonhardening region $[\epsilon^p{}_{B'}, \epsilon^p{}_B]$. When plastic deformation proceeds further to the point C, it expands and translates again to occupy the range $[\epsilon^p{}_C, \epsilon^p{}_{C'}]$. Thus, the nonhardening region eventually occupies the cyclic range of plastic strain, and the hysteresis loop is stabilized. By denoting the size and the center coordinate of the region by 2ρ and α, respectively, it may be represented by

$$g = (\epsilon^p - \alpha)^2 - \rho^2 \leqq 0 \tag{2a}$$

The translation of elastic stress region, R, and the development of its size κ are also shown in Fig. 1. When the plastic strain point ϵ^p moves inside the region $g \leqq 0$, κ remains unchanged, and only the change of R takes place in accordance with the foregoing assumption.

Before prescribing the evolution of the nonhardening region $g \leqq 0$, let us discuss a possible interpretation of it from the microstructural point of view. The region $g \leqq 0$ may be interpreted similarly to the region of recoverable creep hardening in a time-dependent theory [23]. It specifies the temporary softening of metals in creep after stress reversals: As plastic deformation proceeds, dislocations lose thier mobility due to the piling-up of obstacles or the formation of various networks, thus leading to the hardening of metals. When a load reversal occurs, some part of the immobilized dislocations recover the mobility to result in the temporary softening of metals. Since these remobilized dislocations move in the direction opposite the previously immobilized one, they will not form markedly irreversible dislocation structures within a certain region after the load reversal. Thus, it is plausible to assume that the isotropic hardening of metals does not develop within this region of plastic strain. As the reversed straining proceeds, however, the remobilized dislocations are immobilized again gradually; beyond this region, they start to form irreversible dislocation networks.

In the multiaxial state of stress, since the remobilization of dislocation occurs most markedly under reversed loading, this region may have a configuration that is largest in the direction of reversed straining, as shown schematically by the dashed line in Fig. 2. However, if the concern is with stress histories that have no significant deviations from proportional or reversed loading, the nonhardening region may be approximated by the sphere shown in Fig. 2. Thus, equation (2b) is extended to the multiaxial state of stress as follows:

$$g = (2/3)(\epsilon^p_{ij} - \alpha_{ij})(\epsilon^p_{ij} - \alpha_{ij}) - \rho^2 \leqq 0 \tag{2b}$$

where α_{ij} is a symmetric deviator tensor of rank two denoting the center of the region $g \leqq 0$.

As mentioned previously, the nonhardening region expands and translates, only when the plastic strain point is located on the surface $g = 0$ and moves outward. Hence, the evolution of ρ is prescribed as follows:

$$d\rho = c\Gamma d\hat{\epsilon}^p \tag{3}$$

where c is a scalar, possibly a function of the history of plastic deformation, and $d\hat{\epsilon}^p$ and Γ are defined as

$$d\hat{\epsilon}^p = [(2/3)d\epsilon^p_{ij} d\epsilon^p_{ij}]^{1/2} \tag{4}$$

$$\Gamma = \begin{cases} \dfrac{(\epsilon^p_{ij} - \alpha_{ij})d\epsilon^p_{ij}}{(3/2)\rho d\hat{\epsilon}^p}, & g = 0 \quad \text{and} \quad (\partial g/\partial \epsilon^p_{ij})d\epsilon^p_{ij} > 0 \qquad (5a) \\ 0, & g < 0 \quad \text{or} \quad (\partial g/\partial \epsilon^p_{ij})d\epsilon^p_{ij} \leqq 0 \qquad (5b) \end{cases}$$

It is seen that Γ in equation (5a) denotes the cosine of the angle between the plastic strain increment vector and the outward vector normal to the surface $g = 0$. Moreover, a consistency condition $dg = 0$ exists when $g = 0$ and $(\partial g/\partial \epsilon^p_{ij})d\epsilon^p_{ij} > 0$. Hence, if it is assumed that the region $g \leqq 0$ translates in the direction normal to the surface $g = 0$, the evolution of α_{ij} is specified by substituting equation (3) into this condition:

$$d\alpha_{ij} = (1 - c)\Gamma d\hat{\epsilon}^p (\epsilon^p_{ij} - \alpha_{ij})/\rho \tag{6}$$

Although $(\epsilon^p_{ij} - \alpha_{ij})/\rho$ in equations (5a) and (6) becomes indeterminate when $\rho = 0$ $(\epsilon^p_{ij} - \alpha_{ij} = 0)$, it can be evaluated as follows: When $\rho \neq 0$, α_{ij} moves toward ϵ^p_{ij} in the plastic strain space, since $d\alpha_{ij}$ is collinear with $\epsilon^p_{ij} - \alpha_{ij}$ in equation (6). Thus, if the direction of $d\epsilon^p_{ij}$ remains unchanged, $\epsilon^p_{ij} - \alpha_{ij}$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

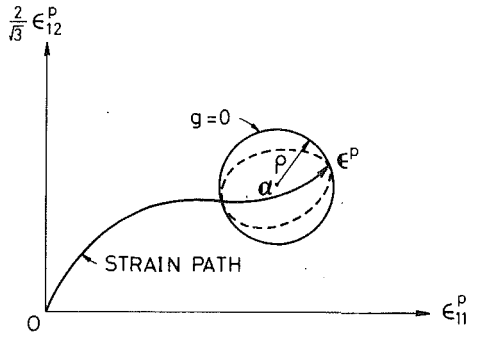


Fig. 2 Nonhardening strain region in multiaxial state of stress

becomes collinear with $d\epsilon_{ij}^p$ as plastic deformation proceeds. In the case of infinitesimally small ρ, this collinearity is realized almost instantaneously as ϵ_{ij}^p moves outward from the surface $g = 0$. Thus, as $\rho \to 0$, $(\epsilon_{ij}^p - \alpha_{ij})/\rho \to d\epsilon_{ij}^p/d\bar{\epsilon}^p$; equations (5) and (6) reduce to $\Gamma = 1$ and $d\alpha_{ij} = (1-c)d\epsilon_{ij}^p$, respectively.

In the particular case of $c = 1/2$, equations (2)–(6) reduce to the evolution equations of the maximum plastic strain range employed by Chaboche et al. [6] (see Section 3).

Incremental Relation Between Stress and Plastic Strain. We are now in a position to specify the changes of isotropic and kinematic hardening variables, q and R_{ij}, in equation (1). From the assumption mentioned previously, the isotropic hardening of materials develops only when the plastic strain point ϵ_{ij}^p is located on the surface $g = 0$ and moves outward. In this sense, this surface is called the isotropic hardening surface. Thus, let us propose an isotropic hardening variable q prescribed as

$$dq = \Gamma d\bar{\epsilon}^p \tag{7a}$$

When $c = 0$ and $\rho(0) = 0$, ρ does not increase from zero. Then, Γ is always equal to one; see the discussion following equation (6). Thus, q becomes identical to the accumulated plastic strain, as seen from equation (7*a*). If $c \neq 0$, by using equation (3), equation (7*a*) can be rewritten as follows:

$$dq = (1/c)d\rho \tag{7b}$$

In this case, the nonhardening region continues to expand with increase in the number of cycles, until it occupies the cyclic range of plastic strain. In other words, $d\rho$ decreases to zero, as ρ increases to half the amplitude of plastic strain. Therefore, it is seen from equation (7*b*) that q has a saturation value depending on the amplitude of plastic strain.

Now let us discuss the change of the kinematic hardening variable R_{ij}. Although Prager [9] and Ziegler [24] prescribed dR_{ij} to be collinear with $d\epsilon_{ij}^p$ and $\sigma_{ij} - R_{ij}$, respectively, their theories are inadequate to describe such phenomena as nonlinear hardening of materials, cyclic relaxation, and cyclic creep. Therefore, more general translation rules of the yield surface have been proposed as mentioned in Section 1. If we consider the theory taking into account the evanescence of strain memory as well as the strain hardening [18], dR_{ij} may be expressed as

$$dR_{ij} = \bar{K}d\epsilon_{ij}^p - K_e d\bar{\epsilon}^p R_{ij}/(R_{kl}R_{kl})^{1/2} \tag{8}$$

where $\bar{K}$ and K_e are, in general, scalar functions of R_{ij}, q, and s_{ij}. The second term of the preceding equation erases the memory of R_{ij}. Thus, this term plays a similar role as recovery terms in constitutive equations of rate-dependent plasticity [18, 25, 26], although the latter are proportional to the time increment instead of $d\bar{\epsilon}^p$.

However, the following modification of the translation rule (8) may be suggested by the previous discussion of the nonhardening region from the microstructural point of view. The variable R_{ij} may represent the back stress acting on pinned or piled-up dislocations (i.e., remobilizable dislocations), as well as the residual stress due to the nonuniform distribution of plastic strain between adjacent crystal grains. As mentioned previously, when the plastic strain point moves inside the region $g \leqq 0$, the rearrangement of remobilized dislocations takes place, which results in the redistribution of the back stress. Therefore, it is suggested that, inside the region $g \leqq 0$, the variable R_{ij} changes more rapidly, while the isotropic hardening variable q remains unchanged (Fig. 1). Let us describe explicitly this rapid change of R_{ij} by expressing the first term of equation (8) as

$$\bar{K}d\epsilon_{ij}^p = Kd\epsilon_{ij}^p + K^* d\epsilon_{ij}^p \tag{9}$$

where the term $K^* d\epsilon_{ij}^p$ represents the aforementioned rapid change of R_{ij}, and hence it vanishes when $\Gamma = 1$.

Let us specify this term by assuming that, in the case of $0 \leqq \Gamma < 1$, the increase of kinematic hardening rate due to the term $K^* d\epsilon_{ij}^p$ is balanced with the suppression of isotropic hardening rate:

$$(\partial f/\partial R_{ij})K^* d\epsilon_{ij}^p = (\partial f/\partial q)(d\bar{\epsilon}^p - dq) \tag{10}$$

Thus, equation (8) becomes

$$dR_{ij} = \left[K + \frac{(\partial f/\partial q)(1-\Gamma)d\bar{\epsilon}^p}{(\partial f/\partial R_{kl})d\epsilon_{kl}^p}\right] d\epsilon_{ij}^p - K_e \frac{R_{ij}d\bar{\epsilon}^p}{(R_{kl}R_{kl})^{1/2}} \tag{11}$$

Now we consider a yield function

$$f = (3/2)(s_{ij} - R_{ij})(s_{ij} - R_{ij}) - \kappa^2(q) \tag{12}$$

and assume the normality rule of $d\epsilon_{ij}^p$ to the yield surface $f = 0$:

$$d\epsilon_{ij}^p = \lambda(s_{ij} - R_{ij}) \tag{13}$$

where λ is a scalar function. Then $(s_{ij} - R_{ij})d\epsilon_{ij}^p = \kappa(q)d\bar{\epsilon}^p$, and equation (11) reduces to

$$dR_{ij} = \left[K + \frac{2}{3}(1-\Gamma)\frac{d\kappa(q)}{dq}\right] d\epsilon_{ij}^p - K_e \frac{R_{ij}d\bar{\epsilon}^p}{(R_{kl}R_{kl})^{1/2}} \tag{14}$$

By using the consistency condition $df = 0$ together with equations (7*a*) and (14), the function λ in equation (13) is specified, and thus $d\epsilon_{ij}^p$ is expressed as

$$\left.\begin{aligned} d\epsilon_{ij}^p &= \frac{3}{2h}\frac{(s_{ij} - R_{ij})(s_{kl} - R_{kl})ds_{kl}}{(s_{mn} - R_{mn})(s_{mn} - R_{mn})} \\ h &= \frac{3}{2}K + \frac{d\kappa(q)}{dq} - \frac{3}{2}K_e \frac{(s_{ij} - R_{ij})R_{ij}}{\kappa(q)(R_{kl}R_{kl})^{1/2}} \end{aligned}\right\} \tag{15}$$

When the plastic strain point reaches the boundary $g = 0$, for example at the points B^* and C^* in Fig. 1, Γ changes discontinuously as seen from equation (5). However, Γ does not appear in the preceding tangential modulus h, and hence the continuity of h is not disturbed by the discontinuous change of Γ.

A constitutive model is provided by equations (15), (12), (7), and (14) together with equations (2)–(6) specifying the nonhardening strain region.

3 Comparisons Between Theoretical and Experimental Results

To verify the present theory, cyclic torsion of thin-walled tubes is analyzed and the theoretical results are compared with those of corresponding experiments. The experiments were performed at room temperature on type 304 stainless steel subjected to a solution heat treatment at 1100°C. Gauge

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

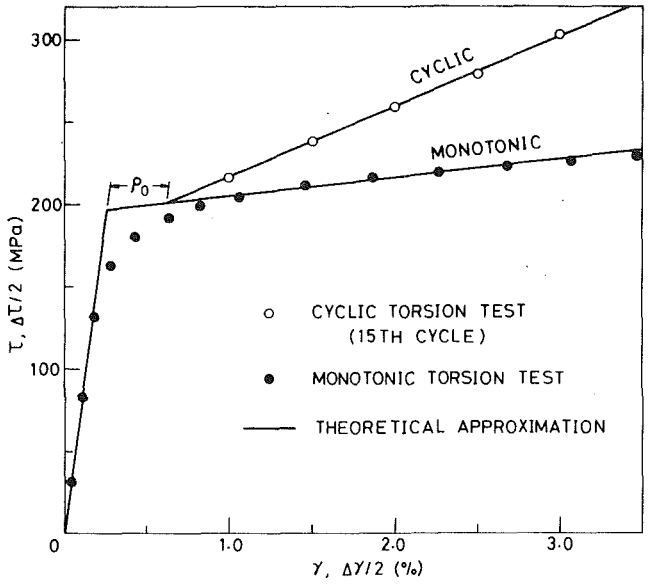


Fig. 3 Stress-strain diagram under monotonic and cyclic torsional loadings

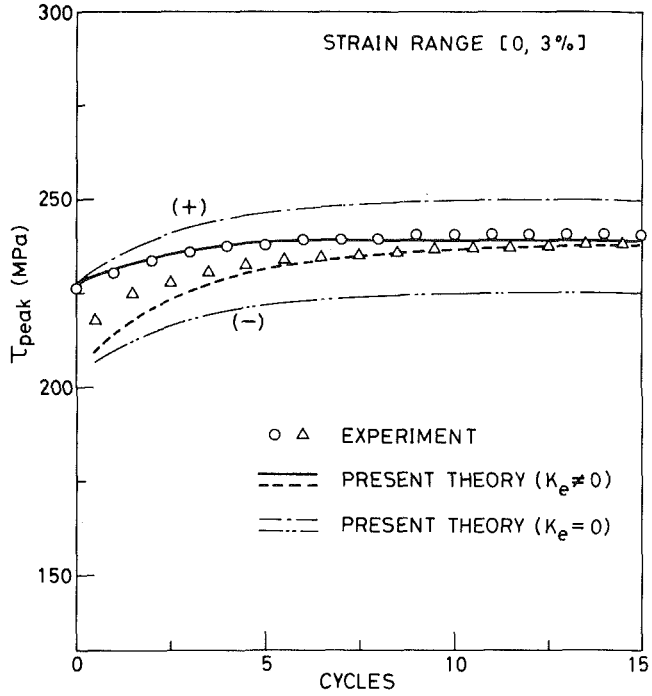


Fig. 4 Saturation and equalization of peak stresses under cycling between unequal and fixed strain limits

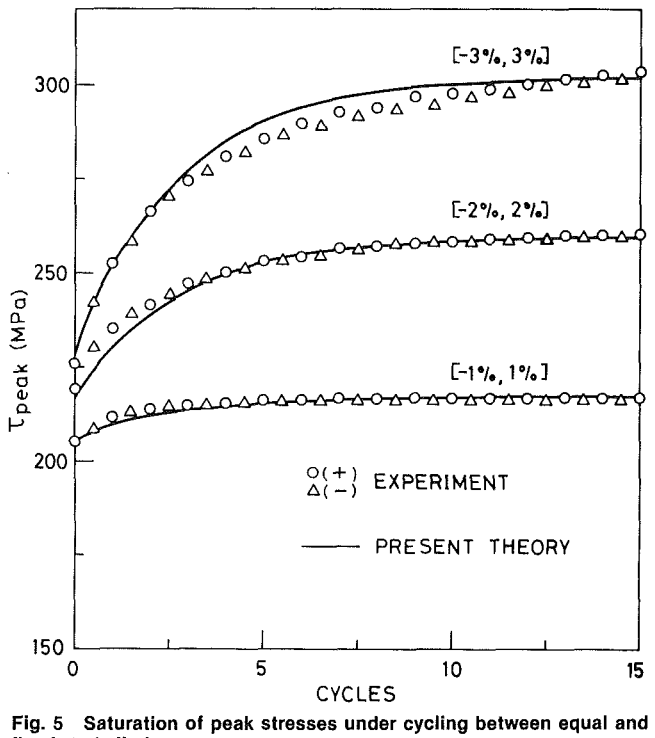


Fig. 5 Saturation of peak stresses under cycling between equal and fixed strain limits

length and both inside and outside diameters of the tubes were 70, 19, and 21 mm, respectively. Cyclic torsional straining was applied under zero axial force at the (engineering) shear strain rate of $\dot{\gamma} = 0.01\ \text{min}^{-1}$.

The experimental results were analyzed within the framework of rate-independent plasticity, though type 304 and 316 stainless steels show rate-dependent deformation behavior at room temperature [27, 28]. A simple model of linear hardening was assumed according to which equations (15), (12), (7*a*), and (14) reduce, within multiplicative factors, to the following equations giving the plastic shear strain increment $d\gamma^p$ under torsional stress τ:

$$d\gamma^p = \{K + L - \text{sgn}[(\tau - R)R]K_e\}^{-1} d\tau \tag{16}$$

$$f = (\tau - R)^2 - (\tau_0 + Lq)^2 \tag{17}$$

$$dq = \Gamma\, |d\gamma^p| \tag{18}$$

$$dR = [K + (1-\Gamma)L]d\gamma^p - \text{sgn}[R]K_e\, |d\gamma^p| \tag{19}$$

where Γ expressing the effect of the nonhardening strain region is specified by equations (2*b*)-(6) as follows:

$$\Gamma = \begin{cases} 1, & g = 0 \quad \text{and} \quad (\partial g/\partial \gamma^p)d\gamma^p > 0 \\ 0, & g < 0 \quad \text{or} \quad (\partial g/\partial \gamma^p)d\gamma^p \leqq 0 \end{cases} \tag{20}$$

$$g = (\gamma^p - \alpha)^2 - \rho^2 \tag{21}$$

$$d\rho = c\Gamma\, |d\gamma^p|, \quad \rho(0) = \rho_0 \tag{22}$$

$$d\alpha = (1-c)\Gamma\, |d\gamma^p|\, \text{sgn}[\gamma^p - \alpha] \tag{23}$$

In the preceding equations, K, L, K_e, τ_0, and c are material constants, and ρ_0 denotes the initial value of ρ. (The initial values of R, q, and α are taken to be zero). In the analysis, moreover, the additive decomposition of strain into elastic and plastic components was assumed; $d\gamma = d\gamma^e + d\gamma^p$, $d\gamma^e = d\tau/G$ (G: elastic shear modulus).

To determine the foregoing material constants, it will be convenient to obtain the relation between the plastic strain amplitude $\Delta\gamma^p$ and the stress amplitude $\Delta\tau$ of a stabilized hysteresis loop. Let us consider the fixed cyclic straining with zero mean strain. In the case of $\Delta\gamma^p > 2\rho_0$, the nonhardening region expands and approaches the cyclic range of plastic strain as the cyclic straining proceeds; when the hysteresis loop is stabilized, $\Gamma_s = 0$ and $\rho_s = \Delta\gamma^p/2$, where the subscript s denotes the stabilized state. Therefore, equations (7*b*) and (19) give, respectively, $q_s = (1/c)(\Delta\gamma^p/2 - \rho_0)$ and $\Delta R_s = (K+L)\Delta\gamma^p$, where $\int \text{sgn}[R]K_e\, |d\gamma^p|$ is nearly equal to zero for half the stabilized cycle. When $\Delta\gamma^p \leqq 2\rho_0$, on the other hand, the region $g \leqq 0$ does not expand (i.e., Γ is always equal to zero, as seen from equation (20)). Thus, we obtain $q_s = 0$ and $\Delta R_s = (K+L)\,\Delta\gamma^p$ from equations (18) and (19). Then,

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

substitution of these q_s and ΔR_s into $\Delta\tau_s = \Delta R_s + 2(\tau_0 + Lq_s)$ gives

$$\frac{\Delta\tau_s}{2} = \begin{cases} 1\tau_0 + (K+L)\rho_0 \\ 1 + \left[K + \left(1 + \dfrac{1}{c}\right)L\right]\left(\dfrac{\Delta\gamma^p}{2} - \rho_0\right), & \Delta\gamma^p > 2\rho_0 \\ \tau_0 + (K+L)\dfrac{\Delta\gamma^p}{2}, & \Delta\gamma^p \leqq 2\rho_0 \end{cases} \tag{24}$$

This equation shows a piecewise linear relation between $\Delta\tau_s$ and $\Delta\gamma^p$. If we use a nonlinear form of $\kappa(q)$ such as $\kappa(q) = (\tau_0{}^n + Lq)^{1/n}$ (n: material constant), they may be related nonlinearly with each other as in [6].

In the determination of material constants, the bilinear representation of monotonic τ-γ diagram gives the plastic tangential modulus $K + L - K_e$ and the initial yield stress τ_0 in equations (16) and (17) (Fig. 3). By approximating $\Delta\tau_s - \Delta\gamma^p$ diagram with equation (24) as shown in Fig. 3, the values of $K + (1 + 1/c)L$ and ρ_0 are determined. Then, if c and K_e are determined, all constants in equations (16)–(23) can be obtained. The values of c and K_e are evaluated to describe, respectively, the saturation of cyclic hardening and the equalization of positive and negative peak stresses, $\tau_{\text{peak}}^{(+)}$ and $\tau_{\text{peak}}^{(-)}$, (Fig. 4). This evaluation of c and K_e presupposes a trial and error numerical procedure employing equations (16)–(23), although c may be determined prior to K_e by tentatively substituting zero for K_e. Thus, the following values were obtained:

$$\left.\begin{array}{lll} K = 925 \text{ MPa}, & L = 265 \text{ MPa}, & K_e = 70 \text{ MPa} \\ \tau_0 = 197 \text{ MPa}, & c = 0.08, & \rho_0 = 0.36 \text{ percent} \end{array}\right\} \tag{25}$$

For the sake of comparison, the theoretical result in the case of $K_e = 0$ and $K = 855$ MPa is also given in Fig. 4. If the memory-erasure term in equation (19) is neglected, the equalization of $\tau_{\text{peak}}^{(+)}$ and $\tau_{\text{peak}}^{(-)}$ (the cyclic relaxation of mean stress) cannot be described.

Figure 5 compares the theoretical predictions and the experimental results on the increase of τ_{peak} under cycling at zero mean strain between fixed strain limits. It is seen that the predictions of equations (16)–(23) express the saturation of cyclic hardening and coincide with the experimental results accurately. The change of stress-strain diagram is shown in the case of $\Delta\gamma/2 = 3.0$ percent in Fig. 6, where the total sum of shear strain, $\int |d\gamma|$, is taken in abscissa. It is observed that the simple model of equations (16)–(23) represents the experimental result within the accuracy of linear hardening.

Let us discuss the case of variable strain limits shown in Fig. 7, where strain limits vary abruptly every 10 cycles in an arbitrary manner as indicated in the figure. Chain lines parallel to the abscissa denote the saturation levels of τ_{peak} in Fig. 5 under cycling between equal and fixed strain limits.

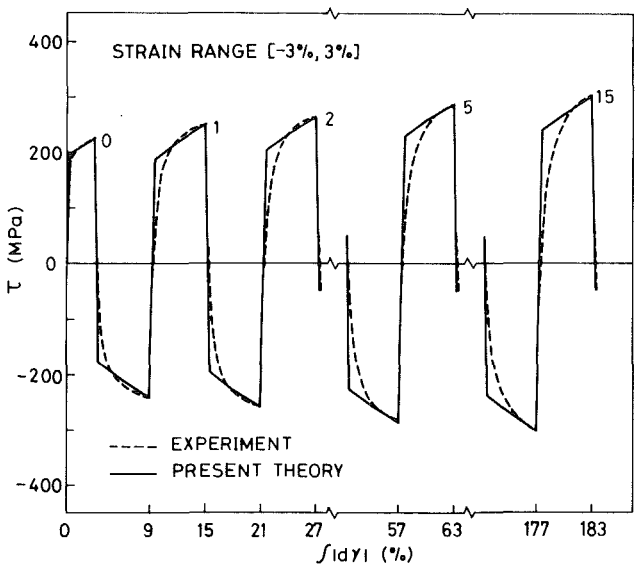


Fig. 6 Change of stress-strain diagram under cycling between equal and fixed strain limits

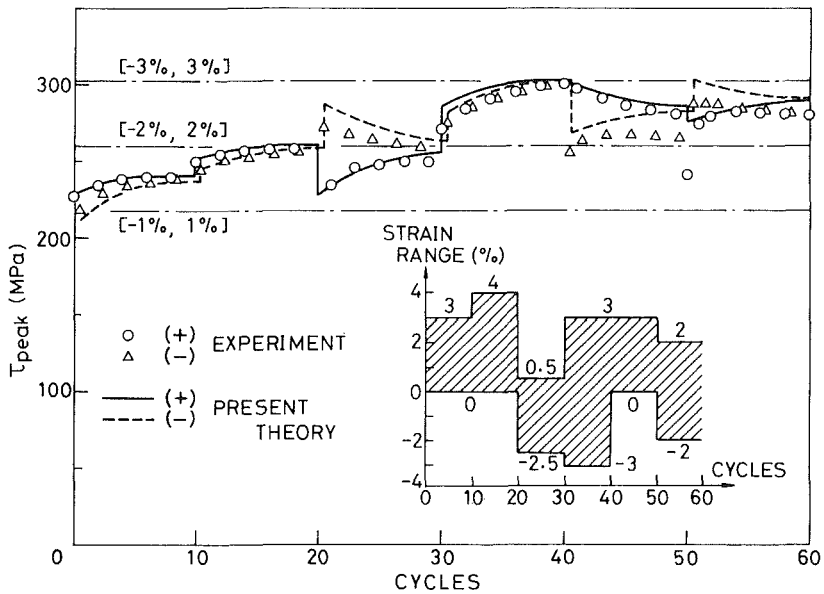


Fig. 7 Change of peak stresses under cycling between variable strain limits

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

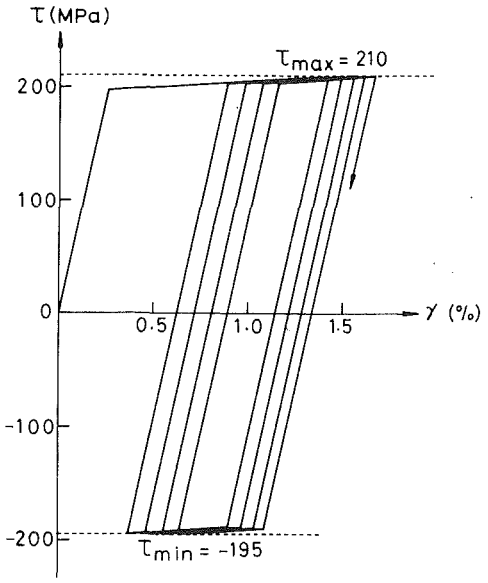


Fig. 8 Accumulation of plastic strain under cyclic loading at nonzero mean stress

As seen from the experimental result, when the current $\Delta\gamma$ is the maximum one which $\Delta\gamma$ has taken so far, the peak stresses increase again after the step-up of $\Delta\gamma$ and then approach asymptotically the saturation level indicated by the chain line corresponding to the current $\Delta\gamma$. It is also seen that the saturation level of τ_{peak} is not influenced by mean strain but only by strain amplitude. When the current $\Delta\gamma$ is less than the maximum one, on the other hand, the peak stresses are not likely to recover to the level of the corresponding chain line at least during the cycles of this experiment. Thus, the cyclic hardening of type 304 stainless steel depends on the maximum amplitude of plastic strain. This experimental trend coincides with those of type 316 stainless steel at room temperature [6] and of type 304 stainless steel at elevated temperature [4], although the experiments in these references were performed at zero mean strain. In Fig. 7, moreover, some differences is generally observed between the magnitudes of $\tau^{(+)}_{peak}$ and $\tau^{(-)}_{peak}$ just after the change of mean strain, even if mean strain is currently zero.

The prediction of the present theory is indicated by solid and dashed lines for $\tau^{(+)}_{peak}$ and $\tau^{(-)}_{peak}$, respectively, in Fig. 7. This prediction expresses the additional increase of peak stresses after the step-up of $\Delta\gamma$, because the nonhardening region expands and translates again to occupy the new cyclic strain range after the increase of $\Delta\gamma$. Also, the present theory always predicts the equalization of $\tau^{(+)}_{peak}$ and $\tau^{(-)}_{peak}$, if there is some difference between them just after the change of mean strain. Thus, the theoretical result expresses the experimental one with fairly good accuracy. As regards the saturation level of τ_{peak} after the decrease of $\Delta\gamma$, however, some discrepancy is observed between the predicted and experimental results. This can be attributed to the fact that cyclic softening is induced very gradually after the decrease of $\Delta\gamma$. Since this cyclic softening cannot be described by the present theory, a memory-erasure term may be necessary in the evolution equation of isotropic hardening.

As mentioned previously, Chaboche et al. [6] prescribed the maximum plastic strain range given by equations (2)–(6) with $c = 1/2$. If elastic strain is neglected, it occupies the strain range [−2.5 percent, 4.0 percent] at the third step in the cyclic deformation history in Fig. 7. Hence, according to their model, the saturation level of τ_{peak} at this step attains that for the constant cyclic strain range [−3.25, 3.25 percent] except for the contribution of kinematic hardening. This prediction does not coincide with the experimental result in Fig. 7, and hence their model is not always valid.

Finally, it is demonstrated in Fig. 8 that the present theory represents the accumulation of plastic strain under cyclic loading at nonzero mean stress. Equations (16)–(23), however, do not express nonlinear stress-strain relations observed just after load reversals, which is important to describe cyclic creep. A more realistic description of cyclic creep may be given, if we consider the dependence of K_e in equation (14) or (19) on R_{ij}.

4 Conclusions

A new constitutive model of cyclic plasticity was presented by specifying a nonhardening strain region inside which the cyclic hardening of a material does not develop. In the simple case of a linear hardening material, the predictions were compared with the experimental results of type 304 stainless steel under cyclic torsional straining. Consequently, the present theory was verified in the case of variable, as well as fixed, strain limits.

However, the present model cannot describe the cyclic softening after the decrease of strain amplitude, and hence a memory-erasure term should be introduced in the evolution equation of isotropic hardening. Specifying a function K_e in equation (14) will make possible the description of nonlinear hardening. These refinements will be pursued in a future study.

Acknowledgments

The author is grateful to Professor S. Murakami of Toyohashi University of Technology for his valuable advice and discussion. The author is also indebted to Professor Y. Ohashi of Nagoya University for his encouragement and use of the testing machine.

References

1 Coffin, L. F., Jr., and Tavernelli, J. F., "The Cyclic Straining and Fatigue of Metals," *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers*, Vol. 215, 1959, pp. 794–806.

2 Feltner, C. E., and Laird, C., "Cyclic Stress-Strain Response of F.C.C. Metals and Alloys — I. Phenomenological Experiments," *Acta Metallurgica*, Vol. 15, 1967, pp. 1621–1632.

3 Krempl, E., "Cyclic Plasticity: Some Properties of Hysteresis Curve of Structural Metals at Room Temperature," ASME *Journal of Basic Engineering*, Vol. 93, 1971, pp. 317–323.

4 Corum, J. M., Greenstreet, W. L., Liu, K. C., Pugh, C. E., and Swindeman, R. W., "Interim Guidelines for Detailed Inelastic Analysis of High-Temperature Reactor System Components," ORNL-5014, 1974.

5 Jaske, C. E., "Consideration of Experimental Techniques Used in Development of Long-Life Fatigue Properties of Pressure Vessel and Piping Alloys," *Characterization of Materials for Service at Elevated Temperature*, Smith, G. V., Ed., ASME, 1978, pp. 129–144.

6 Chaboche, J. L., Dang-Van, K., and Cordier, G., "Modelization of Strain Memory Effect on the Cyclic Hardening of 316 Stainless Steel," *Transactions of the 5th International Conference on Structural Mechanics in Reactor Technology*, Vol. L, North-Holland, 1979, Paper No. L 11/3.

7 Kawashima, K., "Stress-Strain Curve of Pure Aluminum in the Initial Stage of Cyclic Loading," *Journal of the Society of Materials Science, Japan*, Vol. 29, 1980, pp. 351–356 (in Japanese).

8 Hill, R., *The Mathematical Theory of Plasticity*, Oxford, 1950.

9 Prager, W., "A New Method of Analyzing Stresses and Strains in Work-Hardening Plastic Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 23, 1956, pp. 493–496.

10 Dafalias, Y. F., "A Novel Bounding Surface Constitutive Law for the Monotonic and Cyclic Hardening Response of Metals," *Transactions of the 6th International Conference on Structural Mechanics in Reactor Technology*, Vol. L, North-Holland, 1981, Paper No. L 3/4.

11 Bruhns, O., "On the Description of Cyclic Deformation Processes Using a More General Elasto-Plastic Constitutive Law," *Archives of Mechanics*, Vol. 25, 1973, pp. 535–546.

12 Eisenberg, M. A., "A Generalization of Plastic Flow Theory with Application to Cyclic Hardening and Softening Phenomena," ASME *Journal of Engineering Materials and Technology*, Vol. 98, 1976, pp. 221–228.

13 Bodner, S. R., Partom I., and Partom Y., "Uniaxial Cyclic Loading of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Elastic-Viscoplastic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 805–810.

14 Valanis, K. C., "Effect of Prior Deformation on Cyclic Response of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 441–447.

15 Wu, H. C., and Yip, M. C., "Endochronic Description of Cyclic Hardening Behavior for Metallic Materials," ASME *Journal of Engineering Materials and Technology*, Vol. 103, 1981, pp. 212–217.

16 Mróz, Z., Shrivastava, H. P., and Dubey, R. N., "A Non-Linear Hardening Model and Its Application to Cyclic Loading," *Acta Mechanica*, Vol. 25, 1976, pp. 51–61.

17 Eisenberg, M. A., and Phillips, A., "On Nonlinear Kinematic Hardening," *Acta Mechanica*, Vol. 5, 1968, pp. 1–13.

18 Chaboche, J. L., "Viscoplastic Constitutive Equations for the Description of Cyclic and Anisotropic Behavior of Metals," *Bulletin de l'Academie Polonaise des Sciences, Série des Sciences and Techniques*, Vol. 25, 1977, pp. 33–42.

19 Mróz, Z., "An Attempt to Describe the Behavior of Metals Under Cyclic Loads Using a More General Workhardening Model," *Acta Mechanica*, Vol. 7, 1969, pp. 199–212.

20 Eisenberg, M. A., and Phillips, A., "A Theory of Plasticity with Non-Coincident Yield and Loading Surfaces," Vol. 11, 1971, pp. 247–260.

21 Krieg, R. D., "A Practical Two Surface Plasticity Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, 1975, pp. 641–646.

22 Defalias, Y. F., and Popov, E. P., "Plastic Internal Variables Formalism of Cyclic Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 645–651.

23 Murakami, S., and Ohno, N., "A Constitutive Equation of Creep Based on the Concept of a Creep-Hardening Surface," *International Journal of Solids and Structures*, Vol. 18, 1982, pp. 597–609.

24 Ziegler, H., "A Modification of Prager's Hardening Rule," *Quarterly of Applied Mathematics*, Vol. 17, 1959, pp. 55–65.

25 Miller, A., "An Inelastic Constitutive Model for Monotonic, Cyclic, and Creep Deformation," ASME *Journal of Engineering Materials and Technology*, Vol. 98, 1976, pp. 97–113.

26 Kreig, R. D., Swearengen, J. C., and Rohde, R. W., "A Physically Based Internal Variable Model for Rate-Dependent Plasticity," *Inelastic Behavior of Pressure Vessel and Piping Components*, Chang, T. Y., and Krempl, E., Eds., ASME, 1978, pp. 15–28.

27 Kujawski, D., Kallianpur, V., and Krempl, E., "An Experimental Study of Uniaxial Creep, Cyclic Creep and Relaxation of AISI Type 304 Stainless Steel at Room Temperature," *Journal of the Mechanics and Physics of Solids*, Vol. 28, 1980, pp. 129–148.

28 Ellis, J. R., Robinson, D. N., and Pugh, C. E., "Time Dependence in Biaxial Yield of Type 316 Stainless Steel at Room Temperature," *Transactions of the 6th International Conference on Structural Mechanics in Reactor Technology*, Vol. L., North-Holland, 1981, Paper No. L 4/10.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Unified, Self-Consistent Theory for the Plastic-Creep Deformation of Metals

G. J. Weng
Associate Professor,
Department of Mechanics
and Materials Science,
Rutgers University,
New Brunswick, N.J. 08903
Mem. ASME

A unified, self-consistent scheme is formulated to determine the plastic-creep behavior of metals under combined stress. It is pointed out that such a deformation involves the transition from the inhomogeneity to transformation problem in the sense of Eshelby. The plastic deformation is studied by the Berveiller and Zaoui modification of Hill's model. Following plastic deformation the structure of self-consistent relation for subsequent creep is analyzed and found to be independent of prior plastic strains. These self-consistent relations are used in conjunction with one set of unified constitutive equations for slip systems, in which the effect of prior plastic strains on the subsequent creep is considered. This unified, self-consistent scheme is applied to predict the plastic-creep strains of a 304 stainless steel. As compared to the experimental data, the self-consistent scheme is found to consistently provide reasonably accurate estimates for the total inelastic strains, while the predictions by the von Mises theory are seen to be less favorable.

Introduction

When structural metals are subjected to a relatively high stress at elevated temperature, three different modes of deformation are possible: elastic, plastic, and creep. The creep activity in this case is affected by its prior plastic deformation. Traditionally, plastic and creep behavior of metals were studied separately, although many creep theories obviously were generalized from plasticity. The separation of these two topics is understandable under the two extreme conditions: high stress, low temperature (for plasticity) and high temperature, slightly lower stress (for creep). But within certain stress and temperature ranges when both plastic and creep activities are of comparable significance, both types of strains must be considered simultaneously to fully assess the extent of inelastic deformation. To be physically consistent, such a consideration requires a unified treatment. Perhaps due to its inherent complexity, relatively few investigations (e.g., Miller [1] and Pugh and Robinson [2]) have addressed this problem from the unified standpoint. The theories presented are mostly phenomenological, although Miller's work was motivated by the physical mechanisms. These theories are of course useful in their own right. However, the underlying micromechanical process in each grain, which basically links the plastic to creep deformation and which strongly calls for such a unified treatment, and grain interactions, appear not to have been explicitly implemented in these and any other constitutive formulations.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December, 1981; final revision, May, 1982.

The plastic and creep deformation of most structural metals, below half the melting point, is known to be caused primarily by dislocation glide, i.e., crystallographic slip (see, for instance, Dorn and Mote [3], and Ashby and Frost [4]). The "instantaneous" motion is usually identified with plastic deformation and the time-dependent behavior is termed as creep. Since the time-dependent glide takes place following the instantaneous motion, a unified constitutive description is required for the calculation of plastic and creep strains in each grain. Moreover, the heterogeneous stress and strain field throughout the aggregate necessitates the consideration of grain interactions. This perhaps can be most conveniently achieved by means of the self-consistent relations.

In this regard, we recall that, when Eshelby [5, 6] derived the elastic field of an ellipsoidal inclusion in an infinitely extended matrix, two classes of problems were considered: the transformation problem and the inhomogeneity problem. Following the concept of transformation problem, Kröner [7] and Budiansky and Wu [8] formulated the first self-consistent relation for polycrystalline plasticity. Hill [9] later reexamined the self-consistent structure in the context of the inhomogeneity problem and established a more rigorous incremental self-consistent relation. One of its most important implications is that the constraint power of the matrix continues to decrease under monotonic plastic deformation. On the other hand, it was recently pointed out by Weng [10] that in an elastic-creep medium, creep, unlike plastic deformation, is a truly "stress-free" process in the sense of Eshelby, and that Kröner's relation may be used to determine the time-dependent creep behavior of metals.

In this paper we shall use Hill's self-consistent relation to study the plastic response, and then follow the spirit of transformation problem to study the creep behavior of

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

metals. In light of the existence of a prior plastic strain, and with the matrix and inclusion both capable of undergoing plastic deformation, the auxiliary transformation problem will be reexamined to see if Kroner's relation established for an elastic-creep medium is still applicable. More specifically, would the constraint power of the matrix also decrease with an increasing prior plastic strain in the subsequent time-dependent deformation? The constitutive equations proposed in [10] will be extended to describe the unified plastic-creep deformation of a slip system. Finally, the self-consistent theory will be applied to predict the plastic-creep behavior of a stainless steel, and the results compared to experimental data and the popular von Mises theory. Within the framework of small deformation, the elastic anisotropy of the grain and grain rotation will be neglected for simplicity.

Self-Consistent Relations for Plastic-Creep Deformation: A Transition From the Inhomogeneity to Transformation Problem

With the specific objective of determining the plastic-creep behavior under a combined stress we first briefly recapitulate Hill's self-consistent relation for plasticity and Berveiller and Zaoui's modification [11]. The structure of self-consistent relation for subsequent creep will then be analyzed and established.

Hill's Self-Consistent Relation for Elastic-Plastic Deformation. Consider an ellipsoidal inclusion (grain) embedded in an infinitely extended matrix (polycrystalline aggregate). At any generic state the stress and total strain of the grain are denoted by σ and ϵ, respectively, and the counterparts of the aggregate represented by the barred (averaging) quantities $\bar{\sigma}$ and $\bar{\epsilon}$. Under an incremental loading $d\bar{\sigma}$, Hill's self-consistent relation can be written as [9]

$$d\sigma - d\bar{\sigma} = -\mathbf{L}^*(d\epsilon - d\bar{\epsilon}), \tag{1}$$

where $\mathbf{L}^*$ is the constraint tensor of the matrix, related to Eshelby's $\mathbf{S}$, the transformation matrix, through

$$\mathbf{L}^*\mathbf{S} = \mathbf{L}(\mathbf{I} - \mathbf{S}), \tag{2}$$

$\mathbf{L}$ being the tangent moduli of the matrix, and $\mathbf{I}$ the fourth-rank identity tensor. The $\mathbf{S}$ matrix may be determined by solving the transformation problem of an ellipsoidal inclusion in an anisotropic elastic medium with a moduli tensor $\mathbf{L}$, or numerically by Ghahremani's computer program [2]. Using Kneer's solution [13] for the $\mathbf{S}$ matrix with a transversely isotropic $\mathbf{L}$, Hutchinson [14] was able to apply Hill's relation to study the tensile behavior of a polycrystalline aggregate. His results indicate that, except at the early stage of plastic deformation, this scheme leads to a substantially softer plastic response than Kröner's model.

In general, a direct application of Hill's theory is rather involved. However, under a monotonic, proportional loading, if the stress-strain relation of an initially isotropic polycrystal is now expressed in its total form, the secant moduli as measured from the undeformed state will still possess an isotropic symmetry. The values of $\mathbf{S}$ and $\mathbf{L}^*$ then can be derived directly from Eshelby's solution [5] and (2), with the elastic moduli replaced by the corresponding secant moduli. This substitution of incremental formulation by the total form was suggested by Berveiller and Zaoui [11]. With such a substitution, Hill's relation can be modified to

$$\sigma - \bar{\sigma} = -\mathbf{L}_s^*(\epsilon - \bar{\epsilon}), \tag{3}$$

where $\mathbf{L}_s^*$ represents the overall constraint tensor of the matrix in this total deformation, the subscript s referring to the secant moduli. For a spherical inclusion, which will be considered hereafter, the value of $\mathbf{L}_s^*$ can be found by analogy of Hill's $\mathbf{L}^*$ for polycrystalline elasticity [9].

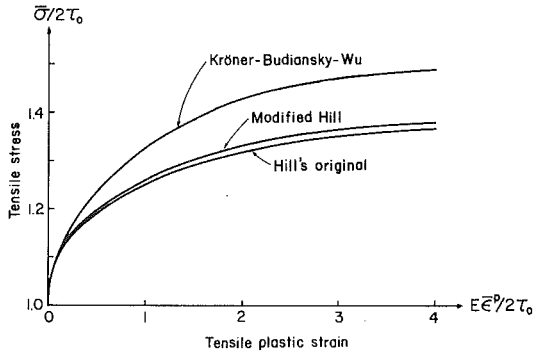


Fig. 1 Comparison of the modified Hill's model with Hill's and Kröner-Budiansky-Wu's models. The crystals were taken to be elastically isotropic, nonhardening with $\nu = 1/3$.

Since the total strain is now the sum of elastic and plastic strains, equation (3) can be rewritten as [11]

$$\sigma = \bar{\sigma} - AG(\epsilon^p - \bar{\epsilon}^p), \tag{4}$$

where A is a plastic constraint factor, G is the elastic shear modulus, and the superscript p refers to the plastic strain. For numerical calculations it was found more convenient to express the secant moduli in terms of the corresponding elastic moduli through two dimensionless parameters:

$$\mathrm{E}_s = p\mathrm{E}, \quad G_s = qG, \tag{5}$$

where E is Young's modulus. Then

$$A = \frac{2q(7 - 5\nu_s)}{q(7 - 5\nu_s) + 2(4 - 5\nu_s)}, \tag{6}$$

where ν_s is the secant Poisson ratio, given by $\nu_s = \mathrm{E}_s/2G_s - 1$. Due to plastic incompressibility only one of the three parameters, p, q, and ν_s, is truly independent; the other two follow from, say

$$\nu_s = 0.5 - (0.5 - \nu)p, \tag{7}$$

$$q = p(1 + \nu)/1 + \nu_s), \tag{8}$$

ν being Poisson's ratio. The ranges of these three parameters lie in: $0 < p \leq 1, 0 < q \leq 1$, and $\nu \leq \nu_s < 0.5$. Since ν_s, as compared to q, is relatively insensitive to plastic deformation, it is evident from (6) that the constraint factor A continues to decrease as plastic deformation persists. Although Hill's relation is greatly simplified, the essence of his scheme, which characterizes a weakening constraint power of a yielded matrix, is still preserved.

To place this modified model in proper perspective we used it to compute the tensile stress-strain relation of a polycrystal. The value of A was constantly adjusted according to (6) for increasing stress and strain state. In this computation the same polycrystal model suggested by Hutchinson [14] was employed, so that his highly accurate calculations based on Hill's and Kröner-Budiansky-Wu's (KBW's) relations could be compared. For the case of nonhardening crystals, the results of the modified and these two models (taken from his Fig. 4) are shown in Fig. 1. It is seen that this modified version provides a far better approximation to Hill's model as compared to the KBW's. Equally good comparisons with Hill's model were also found for the cases of linearly isotropic, work-hardening crystals ($h/\mu = 0.02$ and 0.04 in Hutchinson's Fig. 3).

Self-Consistent Relation for Creep Following Plastic Deformation. For an elastic-creep medium, Kröner's theory was shown [10] to be able to describe the self-consistent relation for the time-dependent creep deformation of polycrystals. His relation was first applied by Brown [15]. In another self-consistent study, Hutchinson [16] has also successfully extended Hill's relation to derive the useful steady creep rate of the aggregate.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

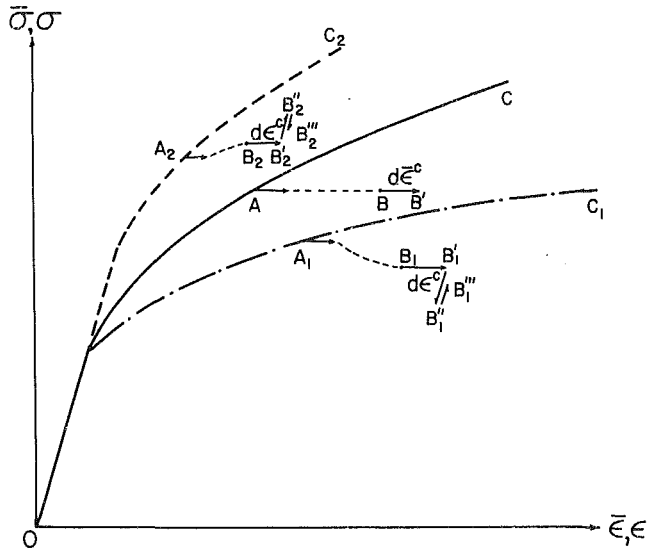


Fig. 2 Schematic representation of incremental creep of the aggregate and of its two representative grains following plastic deformation

When the inclusion and matrix can both deform elastic-plastically under an incremental stress, the self-consistent relation for the subsequent creep deformation qualifies for a reconsideration from the standpoint of transformation problem. Referring to Fig. 2, suppose that the solid line OAC represents the elastic-plastic stress-strain relation of the aggregate, and that the dotted lines OA_1C_1 and OA_2C_2 stand for those of a more favorably and of a less favorably oriented grain, respectively. The state after plastic deformation, but prior to creep, is denoted by points A, A_1, and A_2. Under a constant-stress creep, the stress-strain curve moves horizontally toward the right from A. Its trajectory for a more favorably oriented grain also moves toward the right, but with a decreasing stress, whereas that for a less favorably oriented grain moves with an increasing stress as a function of time. To establish the self-consistent relation for an incremental creep with a prior plastic flow, let us consider the mechanics of transformation problem at an arbitrary state of the creep path, say at B, B_1, and B_2.

Following Eshelby's transformation process, first we allow the inclusion and matrix to creep freely without any constraint, and denote their respective incremental creep strains by $d\epsilon^c$ and $d\bar{\epsilon}^c$. To bring the inclusion back to the matrix compatibly, we apply a uniform strain $-(d\epsilon^c - d\bar{\epsilon}^c)$ to the inclusion. The accumulated creep strains in slip systems serve to raise their flow stresses for plastic deformation; the current stress state therefore lies far below its flow stress at this inelastic strain. For a more favorably oriented grain, characterized by $d\epsilon^c > d\bar{\epsilon}^c$, this straining, as indicated by $B_1'B_1''$ in Fig. 2, is an unloading process, and for less favorably oriented grain, even though this may amount to a further stressing as shown by $B_2'B_2''$, the response is still elastic. Possible elastic-plastic response may occur at the very beginning of creep deformation near A, and only to those least favorably oriented grains. Its effect to subsequent creep of the aggregate is minimal and, for the calculation of entire creep strain history, this straining process may be taken as elastic. Thus, the stress increment in the grain is equal to

$$d\sigma_1 = -\mathcal{L}(d\epsilon^c - d\bar{\epsilon}^c), \tag{9}$$

where $\mathcal{L}$ is its isotropic, elastic moduli tensor. Although the compatibility condition is satisfied at this stage, the equilibrium condition is violated over the grain boundary.

To remove the unwanted layer of surface force we apply a traction $d\mathbf{T}^*$, equal to $d\mathbf{T}^* = -d\sigma_1 \mathbf{m}$ on the welded interface with a unit outward normal $\mathbf{m}$. This traction produces an incremental strain roughly opposite to the early straining process in the inclusion and, because of the constraint of the matrix, its amount, as indicated by $B_1''B_1'''$ or $B_2''B_2'''$, is also less than the previous one. This response is obviously elastic. Moreover, the deformation or the constraint of the matrix, in view of its current stress state and corresponding flow stress, is generally elastic. An elastic-plastic constraint may occur only at the very beginning of creep deformation, and for the overall creep strain history it may again be neglected. In fact, since the creep curve AB lies below OAC, the instantaneous response of the aggregate and of its constituent grains in the consideration of this auxiliary problem is, except near point A, always elastic. Then, according to Eshelby's solution, the stress developed in the grain by the applied surface traction is given by

$$d\sigma_2 = \mathcal{L}\mathcal{S}(d\epsilon^c - d\bar{\epsilon}^c), \tag{10}$$

where $\mathcal{S}$ depends on $\mathcal{L}$.

Thus, following the elastic-plastic response, when the aggregate undergoes a constant-stress creep deformation, the stress increment developed in a grain after an incremental creep may be obtained by the sum of $d\sigma_1$ and $d\sigma_2$. Namely,

$$d\sigma = -\mathcal{L}(\mathbf{I} - \mathcal{S})(d\epsilon^c - d\bar{\epsilon}^c). \tag{11}$$

For a spherical grain it reduces to

$$d\sigma = -2\mu(1-\beta)(d\epsilon^c - d\bar{\epsilon}^c), \tag{12}$$

where $\beta = 2(4-5\nu)/15(1-\nu)$. Equation (12) is the familiar Kröner's relation obtained for an elastic-creep medium [10], and is now also applicable for an elastic-plastic-creep solid under this loading condition. The self-consistent relation for subsequent creep is seen to be independent of the prior plastic strain.

Unified Constitutive Equations

The prediction of plastic-creep strains of a polycrystalline solid is sensitive both to the way in which grain interaction is accounted and to the constitutive description of the inelastic behavior of slip systems. The self-consistent relations therefore must be applied in conjunction with a set of constitutive equations capable of reflecting the essential features of slip activities. Although the precise nature of work hardening on the microstructural scale is still to be fully understood, what has been reported will enable us to establish reasonably useful constitutive equations for slip systems. First, plastic deformation will be considered, and then followed by subsequent creep.

Plastic Deformation. One of the most important features associated with crystallographic slip is the presence of active and latent hardening in slip systems. The active hardening of a slip system is caused by its own dislocation glide and, under monotonic loading, it is a function of its slip strain. Latent hardening, on the other hand, is a result of the slip activity of other slip systems and may be expressed as a function of their slip strains and a latent hardening coefficient, connecting the relative orientation of the considered systems. To describe this behavior, the nature of major hardening mechanisms in metal crystals was recently analyzed by Weng [17]. It was found that some mechanisms, such as dislocation tangles, cell structures, etc., would cause isotropic hardening, whereas dislocation pileups and dislocation rings encircling dispersions would lead to kinematic hardening. A mixed hardening law, which incorporates both isotropic and kinematic hardening and is characterized by the "degree of isotropy in work hardening α," was proposed. This theory will be adopted here.

Since Hill's self-consistent relation has been modified to the total form, the mixed hardening law also needs to be so expressed for numerical calculations. The flow stress-slip strain

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

relation for most metal crystals is nonlinear under single slip; we may write the active hardening of a slip system with the modified Ludwik equation

$$\tau = \tau_0 + h \cdot (\gamma^p)^n, \tag{13}$$

where τ and τ_0 are the current and initial flow stresses, respectively, γ^p is the plastic slip strain of the system, and h and n are two material constants, the value of n lying between 0 and 1. Then, under multiple slips and with the mixed hardening law the current flow stress of a slip system can be expressed as

$$\overset{(i)}{\tau} = \tau_0 + \sum_j \left[\alpha + (1-\alpha)\cos \overset{(i,j)}{\theta} \cos \overset{(i,j)}{\phi}\right] h \cdot (\gamma^p)^n, \tag{14}$$

where the superscript i refers to the ith system, $\overset{(i,j)}{\theta}$ is the angle between the slip directions of the ith and jth systems, and $\overset{(i,j)}{\phi}$ is the angle between their slip plane normals. The quantity in the square bracket is the latent hardening coefficient and the summation over j covers all active slip systems. It is evident from (14) that when $\alpha = 1$ it corresponds to Taylor's isotropic hardening law [18] and when $\alpha = 0$ it represents the kinematic hardening [19]. This hardening law is consistent with the observed Bauschinger effect in single crystals. Hutchinson [14] and Berveiller and Zaoui [11] have used a linear isotropic hardening law in their numerical calculations, which corresponds to the special case: $\alpha = 1$ and $n = 1$.

Creep Following Plastic Deformation. Experimental studies of creep behavior of single crystals at elevated temperature indicate that both steady and transient creep are present (see, for instance, Johnson et al. [20]). The steady creep rate of a slip system depends on its resolved shear stress, and the transient rate decreases because of work hardening. The constitutive equation, which describes the transient creep rate, should therefore be consistent with the work-hardening theory adopted. In the absence of plastic deformation one set of constitutive equations was introduced in conjunction with the mixed hardening law [10]. With a prior plastic deformation, which was caused by the same mechanism and presumably leads to the same type of dislocation structure, the existing plastic slip strains should be added to subsequently generated creep strains, so as to provide a unified account for the transient creep rate. Again using a power law for stress dependency, we write the steady and transient creep rates of the ith slip system, respectively, as

$$\overset{(i)}{\dot{\gamma}_s^c} = \kappa \overset{(i)}{\tau_s^\lambda}, \tag{15}$$

$$\overset{(i)}{\dot{\gamma}_t^c} = \eta \left\{ \overset{(i)}{\zeta \tau_s^\lambda} - \sum_j \left[\alpha + (1-\alpha)\cos \overset{(i,j)}{\theta} \cos \overset{(i,j)}{\phi}\right] (\overset{(j)}{\gamma^p} + \overset{(j)}{\gamma^c}) \right\}, \tag{16}$$

where κ, λ, η, and ζ are four additional single crystal constants, τ_s is the Schmid stress, or the resolved shear stress of the slip system, and γ^c, the sum of the γ_s^c and γ_t^c, is its creep strain. Each of these four constants possesses a unique property: λ describes the stress dependence, κ and ζ, respectively, control the magnitude of steady creep rate and the maximum transient creep of a slip system, and η characterizes how rapidly the transient rate decreases with increasing inelastic strains. It should be noted that the adoption of the power-law function usually results in an easier estimate for these single crystal constants by an inverse method [21], but, as pointed out by Kocks et al. [22], the exponential law generally provides a wider range of stress dependence for the thermally activated glide motion of dislocations.

To assist the determination of creep behavior by von Mises' theory, we write the constitutive equation for the aggregate under tension:

$$\dot{\bar{\epsilon}}_s^c = a \bar{\sigma}^b, \tag{17}$$

$$\dot{\bar{\epsilon}}_t^c = c[d\bar{\sigma}^b - (\bar{\epsilon}^p + \bar{\epsilon}^c)], \tag{18}$$

in parallel to (15) and (16), where a, b, c, and d are polycrystal constants, and the transient creep rate is taken to be nonnegative. No specific hardening law is needed in (18) under pure tension.

Derivation of Plastic-Creep Strains

With the suggested constitutive equations, the plastic-creep strain of the polycrystal can be derived from the properties of its constituent grains self-consistently. Here we establish some basic relations and show how these quantities may be calculated, first for plastic and then for creep deformation.

Plastic Strains. By means of the modified Hill's relation (4), the plastic strain $\bar{\epsilon}_{ij}^p$ at a given $\bar{\sigma}_{ij}$ can be obtained with an iterative process. First, we assume a value for p, lying between 0 and 1, and calculate the corresponding moduli E_s, G_s and ν_s, and the constraint factor A from (5)–(8). Since the total strain consists of elastic and plastic components, we have

$$\bar{\epsilon}_{ij}^p = \frac{1}{2}\left(\frac{1}{G_s} - \frac{1}{G}\right)\bar{\sigma}_{ij} - \left(\frac{\nu_s}{E_s} - \frac{\nu}{E}\right)\delta_{ij}\bar{\sigma}_{kk}, \tag{19}$$

where δ_{ij} is Kronecker's delta. The stress of a grain σ_{ij} is then given in terms of the known $\bar{\sigma}_{ij}$ and $AG\bar{\epsilon}_{ij}^p$, and its unknown plastic strain ϵ_{ij}^p by (4).

The plastic strain of a grain is contributed by the slip strains of its active slip systems; specifically,

$$\epsilon_{ij}^p = \sum_k \overset{(k)}{\nu_{ij}} \overset{(k)}{\gamma^p}, \tag{20}$$

where ν_{ij} is the Schmid factor tensor of a slip system, related to its unit slip direction b_i and slip plane normal n_i through

$$\nu_{ij} = \frac{1}{2}(b_i n_j + b_j n_i). \tag{21}$$

The resolved shear stress of a slip system now can be expressed as

$$\tau_s = \nu_{ij}\sigma_{ij} = \nu_{ij}\left(\bar{\sigma}_{ij} + AG\bar{\epsilon}_{ij}^p - AG\sum_k \overset{(k)}{\nu_{ij}} \overset{(k)}{\gamma^p}\right). \tag{22}$$

For an active slip system its τ_s is equal to its flow stress, given by (14). If there are N active systems ($N \le 5$) in the considered grain, there exist N simultaneous equations for the N unknowns $\overset{(1)}{\gamma^p}, \overset{(2)}{\gamma^p}, \ldots, \overset{(N)}{\gamma^p}$:

$$h\sum_{j=1}^{N}\left[\alpha + (1-\alpha)\cos \overset{(i,j)}{\theta} \cos \overset{(i,j)}{\phi}\right] \cdot (\overset{(j)}{\gamma^p})^n + AG \overset{(i)}{\nu_{kl}} \sum_{j=1}^{N} \overset{(j)}{\nu_{kl}} \overset{(j)}{\gamma^p} = \overset{(i)}{\nu_{kl}} (\bar{\sigma}_{kl} + AG\bar{\epsilon}_{kl}^p) - \tau_0, \tag{23}$$

from their yielding condition, where $i = 1, 2, \ldots, N$. In the numerical calculation one set of potentially active slip systems may be assumed active, and their slip strains calculated by solving these simultaneous equations. With these calculated slip strains the corresponding flow stress of each inactive slip system may be found from (14), and its Schmid stress from

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(22). If the Schmid stress is indeed lower than the flow stress for each of these systems, the assumed set of active systems is the correct one and so is the calculated ϵ_{ij}^p. If any one of these inactive systems does not satisfy this condition, another set of systems should be chosen as active. This process is repeated until the required yield condition is satisfied for all slip systems, and for all grains.

Thus for the assumed p, the plastic strain of each grain is now calculated. With the assumed elastic isotropy for each grain, the plastic strain of the aggregate is given by the average of plastic strain of its constituent grains. Symbolically, we write

$$\bar{\epsilon}_{ij}^p = \{\epsilon_{ij}^p\}, \tag{24}$$

where, and henceforth, the curly bracket represents such an averaging process.

The calculated $\bar{\epsilon}_{ij}^p$ gives rise to a new value for p. If this quantity is equal to the orginally assumed p, or if the calculated $\bar{\epsilon}_{ij}^p$ is equal to (19), this $\bar{\epsilon}_{ij}^p$ is the true solution. If the new p is greater than the old one, a new value even greater than the calculated p should be assumed; conversely, if the calculated p is lower, a lower value should be used. This process is then repeated, until the assumed p and the calculated p are equal, or reasonably close to each other.

Creep Strains After Plastic Deformation. Following plastic deformation the stress and slip strains in each grain are given by the modified Hill's relation, and creep now begins with this initial condition. For each time increment, the steady and transient creep rates of each slip system can be calculated from the constitutive equations (15) and (16), in conjunction with the resolved shear stress and inelastic slip strains of the grain. Since, over a time interval dt, $d\gamma^c = (\dot{\gamma}_s^c + \dot{\gamma}_t^c)dt$, the creep strain increment of a grain is given by

$$d\epsilon_{ij}^c = \sum_k {}^{(k)}\nu_{ij}\, {}^{(k)}d\gamma^c \ . \tag{25}$$

The average of this increment over all grains leads to

$$d\bar{\epsilon}_{ij}^c = \{d\epsilon_{ij}^c\}, \tag{26}$$

for the aggregate.

The incremental creep results in a readjustment of stress state within the aggregate, which can be assessed by Kröner's self-consistent relation (12) for each grain. The new resolved shear stress and inelastic slip strain of each slip system can be readily evaluated, and again by means of (15) and (16) the creep rate of each slip system can be determined for the next time increment. This process is repeated, until the entire creep curve is obtained.

At certain stage of creep deformation if the applied stress is removed, recovery will take place. Within the elastic range the removal of applied stress is accompanied by an instantaneous elastic deformation. The residual stress field and the kinematic hardening mechanisms drive dislocations in more favorably oriented grains to move in reverse, causing a recovery strain. The self-consistent relation for this process is identical to that for creep; its incremental reverse strain, of course, can be calculated in the same step-by-step manner.

von Mises' Theory and Its Associated Flow Rule. For the determination of plastic strains at a given stress, the Hencky-Mises relation will be used for monotonic loading:

$$\bar{\epsilon}_{ij}^p = f(J_2)\bar{\sigma}_{ij}', \tag{27}$$

where J_2 is the second invariant of stress tensor, and $\bar{\sigma}_{ij}'$ is the deviatoric part of $\bar{\sigma}_{ij}$.

The calculation of the subsequent incremental creep on the other hand can be carried out more conveniently by the Prandtl-Reuss relation. In terms of the generalized stress $\bar{\sigma}^*$ and strain $\bar{\epsilon}^*$, the constitutive equations, following (17) and (18) may be written as

$$\dot{\bar{\epsilon}}_s^{*c} = a\bar{\sigma}^{*b}, \tag{28}$$

$$\dot{\bar{\epsilon}}_t^{*c} = c[d\bar{\sigma}^{*b} - \bar{\epsilon}^{*in}], \tag{29}$$

where $\bar{\epsilon}^{*in}$ is the generalized inelastic strain, including plastic and creep strain components. Under combined tension and shear, which will be specifically considered later,

$$\bar{\sigma}^* = (\bar{\sigma}_{11}{}^2 + 3\bar{\sigma}_{12}{}^2)^{1/2}, \tag{30}$$

$$\bar{\epsilon}^{*in2} = \left(\bar{\epsilon}_{11}^{in} + \frac{4}{3}\,\bar{\epsilon}_{12}^{in2}\right)^{1/2}, \tag{31}$$

where $\bar{\epsilon}_{ij}^{in} = \bar{\epsilon}_{ij}^p + \bar{\epsilon}_{ij}^c$. The tensile and shear creep strain rates are related to $\dot{\bar{\epsilon}}^{*c}$ through

$$\dot{\bar{\epsilon}}_{11}^c = \frac{\bar{\sigma}_{11}}{\bar{\sigma}^*}\,\dot{\bar{\epsilon}}^{*c}, \tag{32}$$

$$\dot{\bar{\epsilon}}_{12}^c = \frac{3}{2}\,\frac{\bar{\sigma}_{12}}{\bar{\sigma}^*}\,\dot{\bar{\epsilon}}^{*c}, \tag{33}$$

where $\dot{\bar{\epsilon}}^{*c} = \dot{\bar{\epsilon}}_s^{*c} + \dot{\bar{\epsilon}}_t^{*c}$.

Prediction of Plastic-Creep Strains of a Stainless Steel

To test the applicability of the developed theory we used it to study the plastic-creep behavior of type 304 stainless steel, a principal material being considered in the design of modern power plants. First, following the spirit of inverse method as suggested by Lin et al. [23] for polycrystalline plasticity and by Weng [21] for creep, the seven single-crystal constants – two in plasticity, four in creep, and one shared by both – were determined by the simulation of two tensile experimental data, both involving plastic and creep deformation and one followed by recovery. The obtained constants were then used

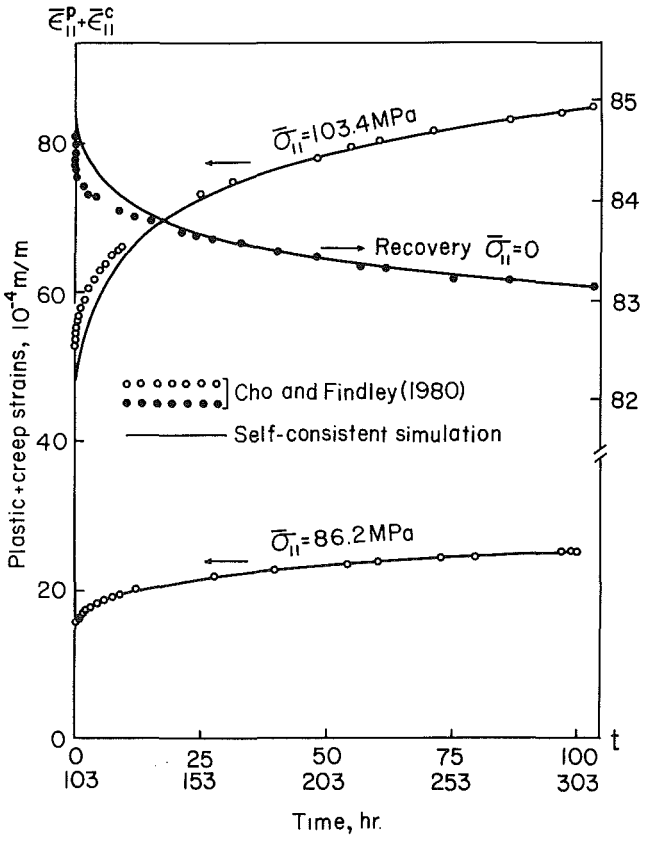


Fig. 3 Derivation of single-crystal constants by the simulation of tensile plastic-creep deformation at two stress levels (upper time scale), one followed by recovery (lower time scale)

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

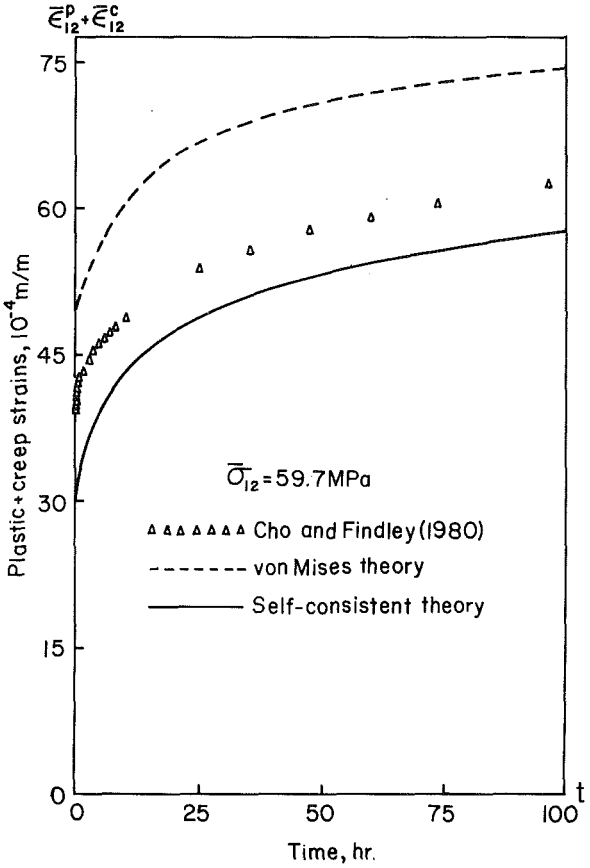


Fig. 4 Predictions of plastic-creep strains under pure shear

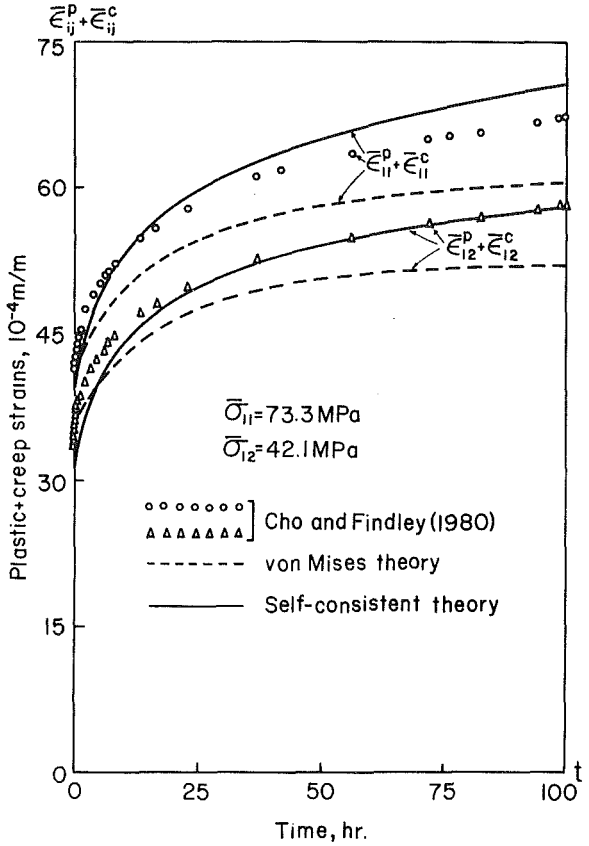


Fig. 5 Predictions of plastic-creep strains under combined tension and shear

to determine the plastic-creep strains of the same material under other loading conditions. Cho and Findley's [24] carefully measured experimental data were used both for derivation of crystal constants and for comparison with predictions.

Most steels have body-centered-cubic (b.b.c.) crystal lattice. Type 304 stainless steel, however, is austenitic and possesses a faced-centered-cubic (f.c.c.) structure, which is characterized by four {111} slip planes and three <110> slip directions on each plane. If the slip systems of a b.c.c. crystal are taken to be {110}<111>, its plastic-creep strains are identical to those of an f.c.c. crystal, although an opposite sense of rotation is expected. In the numerical calculation, the polycrystal model consists of 125 equally sized spherical grains, whose orientations were so chosen to simulate the three-dimensional isotropy. These orientations were generated by a sequence of rotation of a basic crystal about the loading axes, the basic crystal being so originally aligned. Due to the cubic symmetry, the rotation only needs to cover the range from 0 to $\pi/2$. We divided it into five stations, starting from 9 deg and ending at 81 deg, each 18 deg apart. This would avoid the eight-fold symmetry under tension, which in reality is not likely to occur to 1/125 of entire grains. The rotation took place under the sequence of x_3-x_2-x_1 axes, so that under pure tension $\bar{\sigma}_{11}$, only 25 grains need to be considered. As compared to the 75-grain model used previously [10], the present one offers a better isotropy for plastic deformation, although for creep, in which every slip system is active, these two models differ only slightly. The maximum anisotropy, as measured by the generated inelastic strain components, was found to be less than 5 percent for the present model.

Cho and Findley's experiments on plastic-creep deformation were conducted at 593°C, at which the initial tensile yield stress is 57.21 MPa, and the Young's and shear moduli are: $E=14.8\times10^4$ MPa and $G=5.56\times10^4$ MPa, respectively [24]. The initial flow stress of the slip system, with the assumed elastic isotropy, is taken as $\tau_0=28.61$ MPa. The two tensile tests were conducted at $\bar{\sigma}_{11}=103.4$ MPa and 86.2 MPa; the corresponding plastic and creep strains were reproduced in Fig. 3. The recovery strain following the removal of $\bar{\sigma}_{11}=103.4$ MPa after 103 hr of creep is also plotted in the same figure. It is usually difficult to identify the exact amount of plastic strain from such a figure but, because a greater plastic strain would result in a lower subsequent creep rate, and vice versa, a pin-point accuracy is not required for it if one's objective is on the *total* inelastic strain generated. The values of plastic strain are listed in Table 1 of their paper; under these two stresses they are 48.06×10^{-4} m/m and 14.89×10^{-4} m/m, respectively. For the calculation of plastic strains in the self-consistent scheme, the system of nonlinear equations (23) was solved by the Newton-Bailey method, in conjunction with the quadratic Taylor polynomials and Guass-Jordon's direct elimination. Based on their experimental results the single-crystal constants were found to be: $h=2.19\times10^2$, $n=0.72$, $\kappa=1.50\times10^{-17}$, $\lambda=7.41$, $\eta=4.20\times10^{-2}$, $\zeta=2.66\times10^{-14}$, and $\alpha=0.276$, where stress, strain, and time are in the units of MPa, m/m and hr, in turn. With these constants the calculated plastic, creep, and recovery strains by the self-consistent method are also depicted in Fig. 3 as solid lines; the test data are seen to be reasonably well simulated. Further, with the stated plastic strains, these two tensile creep curves could also be well described by (17) and (18), with $a=3.37\times10^{-19}$, $b=6.67$, $c=4.40\times10^{-2}$, and $d=3.05\times10^{-16}$, where stress, strain, and time, are in the same units as just mentioned. In these simulations and the following numerical calculations the time increments were chosen as: for creep, $\Delta t=0.025$ if $t<0.1$; $\Delta t=0.15$ if $0.1\le t<4$; $\Delta t=0.4$ if $4\le t<10$; $\Delta t=0.5$ if $10\le t<60$; $\Delta t=1$ if $60\le t<103$; and for recovery, $\Delta t=1$ if

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Plastic and creep strain components at the end of 100 hr, 10^{-4} m/m

Method	$\bar{\sigma}_{12} = 59.7$ MPa		$\bar{\sigma}_{11} = 73.3$ MPa & $\bar{\sigma}_{12} = 42.1$ MPa			
	$\bar{\epsilon}^p_{12}$	$\bar{\epsilon}^p_{12} + \bar{\epsilon}^c_{12}$	$\bar{\epsilon}^p_{11}$	$\bar{\epsilon}^p_{12}$	$\bar{\epsilon}^p_{11} + \bar{\epsilon}^c_{11}$	$\bar{\epsilon}^p_{12} + \bar{\epsilon}^c_{12}$
Test data	36.14	62.61	38.47	31.86	67.33	58.51
von Mises	49.93	74.09	40.87	35.21	60.61	52.22
Self-consistent	29.82	56.72	38.63	30.87	71.16	58.49

$103 \leq t < 140$; $\Delta t = 2$ if $140 \leq t < 180$; $\Delta t = 5$ if $t \geq 180$, in the unit of hr.

The same material was also tested under pure shear, and under combined tension and shear. We used the foregoing single-crystal constants and the self-consistent relations to calculate the plastic and creep strains for these loading conditions. The results were found to be in generally good agreement with the experimental data. Two selected cases, one under pure shear and another under combined tension and shear, both having the same J_2 as $\bar{\sigma}_{11} = 103.4$ MPa, are shown in Figs. 4 and 5, respectively. The pure shear was conducted at $\bar{\sigma}_{12} = 59.7$ MPa, and the combined stress at $\bar{\sigma}_{11} = 73.3$ MPa and $\bar{\sigma}_{12} = 42.1$ MPa, each lasting for 100 hr. The corresponding results by the self-consistent scheme and by von Mises' theory are also shown in the same figures, as solid and dashed lines, respectively. The plastic and creep strain components by experiments and by these two theories are listed in Table 1. The experimental data for plastic strain were again taken from Table 1 of Cho and Findley [24] and the sum of plastic and creep strains were taken from their Figs. 1 and 2, minus the elastic strain components listed in the same table. It is seen that at the end of 100 hr, the predicted inelastic strain components by the self-consistent scheme are fairly close to the experimental data in both cases. On the other hand, von Mises' theory overestimates the inelastic strain by 18 percent under pure shear, and underestimates the tensile and shear components by 10 and 11 percent, respectively.

As compared to the experimental results and von Mises' theory, the self-consistent scheme is seen to consistently provide reasonably accurate estimates for the total inelastic strains under a combined stress.

Acknowledgment

The financial supports of the National Science Foundation, Solid Mechanics Program, through Grant CME-80-19546 for the study of metal plasticity, and of the U.S. Department of Energy, through Contract DE-AS05-81ER10976 for the study of creep, are gratefully acknowledged.

References

1 Miller, A., "An Inelastic Constitutive Model for Monotonic, Cyclic, and Creep Deformation: Part 1—Equations Development and Analytical Procedures," ASME *Journal of Engineering Materials and Technology*, Vol. 98, 1976, pp. 97–105.

2 Pugh, C. E., and Robinson, D. N., "Some Trends in Constitutive Equation Model Development for High-Temperature Behavior of Fast-Reactor Structural Alloys," *Nuclear Engineering and Design*, Vol. 48, 1978, pp. 269–276.

3 Dorn, J. E., and Mote, J. D., "On the Plastic Behavior of Polycrystalline Aggregate," *Material Science Research*, Vol. 1, 1963, pp. 12–56.

4 Ashby, M. F., and Frost, H. J., "The Kinetics of Inelastic Deformation Above 0°K," *Constitutive Equations in Plasticity*, Argon, A. S., ed., M.I.T. Press, Cambridge, Mass., 1975, pp. 117–147.

5 Eshelby, J. D., "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proceedings of the Royal Society*, London, Series A, Vol. 241, 1957, pp. 376–396.

6 Eshelby, J. D., "Elastic Inclusions and Inhomogeneities," *Progress in Solid Mechanism*, Vol. 2, Sneddon, I. N., and Hill, R., eds., North-Holland, Amsterdam, 1961, pp. 89–140.

7 Kröner, E., "Zur Plastischen Verformung des Vielkristalls," *Acta Metallurgica*, Vol. 9, 1961, pp. 155–161.

8 Budiansky, B., and Wu, T. T., "Theoretical Prediction of Plastic Strains of Polycrystals," *Proceedings, 4th U.S. National Congress of Applied Mechanics*, 1962, pp. 1175–1185.

9 Hill, R., "Continuum Micro-Mechanics of Elastoplastic Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 13, 1965, pp. 89–101.

10 Weng, G. J., "Self-Consistent Determination of Time-Dependent Behavior of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 41–46.

11 Berveiller, M., and Zaoui, A., "An Extension of the Self-Consistent Scheme to Plastically-Flowing Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 26, 1979, pp. 325–344.

12 Ghahremani, F., "Numerical Evaluation of the Stresses and Strains in Ellipsoidal Inclusions in an Anisotropic Elastic Material," *Mechanics Research Communications*, Vol. 4, 1977, pp. 89–91.

13 Kneer, G., Über die Berechnung der Elastizitatsmoduln Vielkristalliner Aggregate mit Textur," *Physica Status Solidi*, Vol. 9, 1965, pp. 825–838.

14 Hutchinson, J. W., "Elastic-Plastic Behavior of Polycrystalline Metals and Composites," *Proceedings of the Royal Society*, London, Series A, Vol. 319, 1970, pp. 247–272.

15 Brown, G. M., "A Self-Consistent Polycrystalline Model for Creep Under Combined Stress States," *Journal of the Mechanics and Physics of Solids*, Vol. 18, 1970, pp. 367–381.

16 Hutchinson, J. W., "Bounds and Self-Consistent Estimates for Creep of Polycrystalline Materials," *Proceedings of the Royal Society*, London, Series A, Vol. 348, 1976, pp. 101–127.

17 Weng, G. J., "Dislocation Theories of Work Hardening and Yield Surfaces of Single Crystals," *Acta Mechanica*, Vol. 37, 1980, pp. 217–230.

18 Taylor, G. I., "Plastic Strain in Metals," *Journal of the Institute of Metals*, Vol. 62-1, 1938, pp. 307–324.

19 Weng, G. J., "Kinematic Hardening Rule in Single Crystals," *International Journal of Solids and Structures*, Vol. 15, 1979, pp. 861–870.

20 Johnson, R. D., Shober, F. R., and Schope, A. D., "The Creep of Single Crystals of Aluminum," National Advisory Committee for Aeronautics, Technical Note 2945, 1953.

21 Weng, G. J., "A Physically Consistent Method for the Prediction of Creep Behavior of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 800–804.

22 Kocks, U. F., Argon, A. S., and Ashby, M. F., "Thermodynamics and Kinetics of Slip," *Progress in Materials Science*, Vol. 19, Chalmers, B., Christian, J. W., and Massalski, T. B., eds., Pergamon Press, 1975.

23 Lin, T. H., Ito, Y. M., and Yu, C. L., "A New Slip Theory of Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 587–592.

24 Cho, U. W., and Findley, W. N., "Creep and Creep Recovery of 304 Stainless Steel Under Combined Stress with a Representation by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 755–761.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Y. Weitsman
Professor,
Mechanics and Materials Center.
Mem. ASME

B. D. Harper
Graduate Student.

Texas A&M University,
College Station, Texas 77843

Optimal Cooling of Cross-Ply Composite Laminates and Adhesive Joints

This paper concerns the optimal cooling of symmetric, balanced, cross-ply composite laminates and adhesive joints so as to minimize the residual thermal stresses upon termination of the cool-down process. The computations are based on a recently developed analytical scheme and employ up-to-date data on graphite/epoxy laminas. The calculations consider the thermoviscoelastic response of the polymeric resins and incorporate the temperature dependence of the coefficients of thermal expansion. It is shown that the viscoelastic behavior may contribute to a significant reduction of the residual stresses.

Introduction

This paper presents results based on a recently developed optimization scheme [1][1] and data on viscoelastic response of graphite/epoxy composites [3] whereby specific time-temperature paths are selected during the cool-down stage so as to minimize the residual thermal stresses in cross-ply laminates and in adhesively bonded aluminum/epoxy joints.

The cross-ply laminates under consideration consist of a symmetric and balanced lay-up of 0 and 90 deg plies. When such laminates are cooled from the presumed stress-free cure temperature to room temperature, the stiffer fibers within each ply restrain the transverse shrinkage of the adjacent, perpendicularly directed plies and these mutual geometric constraints introduce substantial residual stresses in the laminate.

Consider a finite cool-down time. During this time interval the temperature introduces two counteracting effects. On one hand, since larger temperature drops induce larger residual stresses it appears that sharp and early temperature drops will generate significant stresses that may undergo relaxation. On the other hand, the relaxation process is retarded at lower temperatures. The optimal time-temperature path achieves the best interplay between the aforementioned competing effects so as to minimize the residual thermal stress at the termination of the cool-down phase.

Elastic Analysis

Consider a symmetric, balanced, cross-ply composite laminate, consisting of equal number of 0 and 90 deg plies exposed to fluctuating ambient temperature $T(t)$. In view of the small thicknesses of the composite laminates employed in practice it is permissible to ignore the process of thermal diffusion and employ the approximation that $T = T(t)$ throughout the laminate. A similar circumstance also holds for adhesive joints.

Denoting by subscripts L and T the longitudinal (parallel to fibers) and transverse (perpendicular to fibers) directions, respectively, we have the following stress-strain relations in each individual lamina [4, 5]

$$\epsilon_L - \alpha_L \Delta T = \frac{\sigma_L}{E_L} - \frac{\nu_{TL}}{E_T}\sigma_T$$
$$\epsilon_T - \alpha_T \Delta T = -\frac{\nu_{LT}}{E_L}\sigma_L + \frac{\sigma_T}{E_T} \tag{1}$$

In (1) ϵ and σ designate strain and stress, while α, E, and ν denote coefficient of thermal expansion, modulus, and Poisson's ratio, respectively. ΔT represents temperature difference.

In view of the particular lay-up under consideration we have

$$\sigma_T = -\sigma_L = \sigma, \quad \epsilon_T = \epsilon_L = \epsilon$$

and straightforward manipulations yield

$$\sigma = -\frac{E_T(\alpha_T - \alpha_L)\Delta T}{\left(1+\nu_{TL}\right)\left(1+\frac{1+\nu_{LT}}{1+\nu_{TL}}\frac{E_T}{E_L}\right)} \tag{2}$$

For typical graphite/epoxy laminae we have $E_L \sim 20E_T$ (whereby $\nu_{LT} \sim 20\nu_{TL}$). With $\nu_{LT} \simeq 0.28$ we obtain

$$\sigma = -rE_T(\alpha_T - \alpha_L)\Delta T \tag{3}$$

Typically $r \approx 0.90$–0.93.

Alternately, consider two adherend plates, each of thickness h, bonded together by an adhesive layer of thickness

[1] See also [2] for treatment of the same problem.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April, 1981; final revision, November, 1981. Paper No. 82-WA/APM-24.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

a. Let the joint undergo a spatially uniform temperature excursion ΔT. Elementary calculations then yield

$$\sigma_e = \frac{E_e}{1-\nu_e} \frac{(\alpha_e - \alpha_A)\Delta T}{1 - \frac{1-\nu_A}{1-\nu_e}\frac{E_e}{E_A}\frac{a}{2h}} \tag{4}$$

In (4) subscripts "A" and "e" refer to the adherend and adhesive, respectively. For typical joints $E_A/E_e = 50$ and $h/a = 100$. Consequently (4) can be approximated by

$$\sigma_e \cong \frac{E_e}{1-\nu_e}(\alpha_e - \alpha_A)\Delta T \tag{5}$$

Viscoelastic Formulation

Data on graphite/epoxy laminas [3] and on "neat" epoxy resins [6] indicate that E_T and E_e are time and temperature dependent.[2] The time dependence can be approximated by the form

$$E = E_0 (t + t_0)^{-q} \tag{6}$$

while the temperature dependence can be related by means of a shift-factor function $a(T)$, with $a(T_R) = 1$, where customarily T_R is room temperature (293°K). Consequently (6) is extended to read

$$E = E_0 \left\{ t_0 + \left[t/a(T) \right] \right\}^{-q} \tag{7}$$

Available data [3, 6] indicate that $a(T)$ can be expressed by

$$a(T) = \exp\left(-\frac{T}{A} + B \right) \tag{8}$$

Employing the superposition integral, equations (3) and (5) yield [7]

$$\sigma(t) = \alpha E_0 \int_0^t \left[\int_\tau^t \frac{ds}{a(T(s))} + t_0 \right]^{-q} \frac{d\Delta T(\tau)}{d\tau} d\tau \tag{9}$$

In (9) and the sequel, α represents $(\alpha_T - \alpha_L)$ or $(\alpha_e - \alpha_A)$ and E_0 absorbs the factors r or $1/(1-\nu_e)$, respectively. It should be noted that in the viscoelastic case the factor r in equation (3) becomes a complicated function of time and its incorporation into the subsequent analysis requires cumbersome details. However, in all the practical cases which we considered $r(t)$ varied at most by 1/2 percent about its average value over the time span of interest. Consequently we employed the approximation $r(t) = r = \text{const.}$ in our work, without sacrificing accuracy. Data on graphite/epoxy [3] indicate that α_T and α_L vary significantly with temperature. This aspect of the material response will be considered later.

Optimal Cooling Paths

Consider now the case of cooling from the initial cure temperature T_I to room temperature T_R in a given time interval t_f. It has been shown [1, 2] that if an optimal path exists then it is possible to obtain a minimal value for $\sigma(t)$ in (9). In such a case this minimal value of $\sigma(t)$ is obtained by following a path $T(t)$ that includes an abrupt initial drop from T_I to T_0. This initial drop was given by solving the transcendental equation [1]

$$T_0 - T_I = \frac{a(T_0)}{a'(T_0)} \tag{10}$$

In the open interval $0 < t < t_f$ the optimal path $T(t)$ was shown to be smooth and continuous, where it is governed by the nonlinear integrodifferential equation

$$\frac{dT}{dt} = -\frac{E''(t,t_f)}{E'(t,t_f)} \frac{a'(T(t))}{a(T(t))a''(T(t))} \tag{11}$$

[2] These data are corroborated by many other sources and publications, which are not listed here.

In (10) and (11) primes indicate derivatives with respect to the argument. Also

$$E(t_1, t_2) = E\left[\int_{t_1}^{t_2} \frac{ds}{a(T(s))} \right]$$

With given $E(t)$, $a(T)$, and T_I, the value of T_0 is determined by (10) and the optimal temperature path can be constructed by iteration through the employment of (11). In most cases the final prescribed temperature T_R is reached through a second abrupt step at time t_f.

A particular simplification occurs for a "power law" response, as expressed in (6) and, in addition, when the shift-factor function is given by (8). In this case it was shown [1] that the optimal path is given by

$$T = \frac{1}{C} \ln \frac{L}{k + Ct} \tag{12}$$

where, in (12)

$$C = \frac{q}{(q+1)A} \quad \text{and} \quad L = K \exp(CT_0)$$

The term K is determined as the root of the transcendental equation

$$Ke^{CT}0 = \left[\frac{e^B t_0 (K + Ct_f)^{\frac{1}{q}}}{(q+1)A} \right]^{\frac{q}{q+1}} \tag{13}$$

Under the simplifications that occur in the present circumstances the iterative process alluded to in [1] reduces to the search for the root K of equation (13).

In view of (8), equation (10) yields

$$T_0 = T_I - A \tag{14}$$

With the optimal temperature-path prescribed by (12) the stresses are obtained by direct substitution of (12) into (9). After several manipulations we get

$$\sigma(t) = \alpha E_0 \left\{ \int_0^t Q(\tau, t)\, d\tau + K(T_I - T_0) Q(0,t) \right\} \tag{15}$$

where

$$Q(\tau,t) = \frac{t_0^{-q} \left\{ \left(\frac{K+Ct_f}{K+Ct} \right)^{\frac{1}{q}} \left[\left(\frac{K+Ct}{K+C\tau} \right)^{\frac{1}{q}} - 1 \right] + 1 \right\}}{K + C\tau}$$

$\sigma(t_f^-)$ is obtained by setting $t = t_f^-$ in (15). At $t = t_f^+$ we get

$$\sigma(t_f^+) = \sigma(t_f^-) + \alpha E_0 t_0^{-q} (T_f - T_R) \tag{16}$$

Temperature-Dependent Coefficients of Thermal Expansion

Recent data on graphite/epoxy composites indicate that the quantity $(\alpha_T - \alpha_L)$ increases monotonically with temperature. Consequently (9) must be modified to read

$$\sigma(t) = E_0 \int_0^t \left[\int_\tau^t \frac{ds}{a(T(s))} + t_0 \right]^{-q} \frac{d}{d\tau} \left[\alpha(T(\tau)) \Delta T(\tau) \right] d\tau \tag{17}$$

where $\alpha(T(t)) = \alpha_T(T(t)) - \alpha_L(T(t))$.

Consider the thermal strain

$$\epsilon_{th}(t) = \alpha(T(t))\, \Delta T(t) \tag{18}$$

In view of the one-to-one correspondence between ϵ_{th} and T we have

$$a(T(t)) = \beta(\epsilon_{th}(t)) \tag{19}$$

Therefore (17) can be rewritten as follows

$$\sigma(t) = E_0 \int_0^t \left[\int_\tau^t \frac{ds}{\beta(\epsilon_{th}(s))} + t_0 \right]^{-q} \frac{d}{d\tau} (\epsilon_{th}(\tau))\, d\tau \tag{20}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Material parameters

Material system		E_0 (KN/m^2)	r	q	t_0 (min)	α (cm/cm°K)	A (°K)	B
graphite/epoxy	AS/3502	$10^{7.055}$	0.9	0.00775	1	$\alpha_L = -0.217\times10^{-6}$ $\alpha_T = 24.73\times10^{-6}$ (at $T=372$°K)	6.258	46.81
	T300/5208	$10^{7.014}$	0.93	0.00717	1	$\alpha_L = -0.593\times10^{-6}$ $\alpha_T = 25.65\times10^{-6}$ (at $T=372$°K)	6.792	43.14
aluminum epoxy		$10^{6.54}$	1	0.036	0.5	$\alpha_e = 5\times10^{-5}$ $\alpha_A = 2.4\times10^{-5}$	7.91	37.05

Table 2 The temperature dependence of the thermal expansion coefficients. $\alpha(T) = \alpha_0 + \alpha_1 T$.

Material system	α_L (cm/cm/°K)		α_T (cm/cm/°K)	
	α_0 (cm/cm/°K)	α_1 (cm/cm/°K^2)	α_0 (cm/cm/°K)	α_1 (cm/cm/°K^2)
AS/3502	-0.484×10^{-6}	0.72×10^{-9}	9.5×10^{-6}	0.41×10^{-7}
T300/5208	-0.418×10^{-6}	-0.47×10^{-9}	11.9×10^{-6}	0.37×10^{-7}

Note that (20) is analogous to (9). Accordingly, (10) and (11) yield the following expressions for the optimal path

$$\epsilon_{th}^0 - \epsilon_{th}^f = \frac{\beta(\epsilon_{th}^\circ)}{\beta'(\epsilon_{th}^\circ)} \tag{21}$$

and

$$\frac{d\epsilon_{th}(t)}{dt} = -\frac{E''(t,t_f)}{E'(t,t_f)} \frac{\beta'(\epsilon_{th}(t))}{\beta(\epsilon_{th}(t))\beta''(\epsilon_{th}(t))} \tag{22}$$

In (22), $E(t_1,t_2) = E\left[\int_{t_1}^{t_2} \frac{ds}{\beta(\epsilon_{th}(s))}\right]$.

Equations (21) and (22) can be expressed in terms of $a(T)$ and T through the reintroduction of (18) and (19). We obtain

$$T_0 - T_I = \frac{\alpha(T_0)a(T_0)}{\alpha(T_0)a'(T_0) - \alpha'(T_0)a(T_0)} \tag{23}$$

$$\frac{dT}{dt} = -\frac{E''(t,t_f)}{E'(t,t_f)} \frac{a'}{a\left[a'' - a'\frac{2\alpha' + \alpha''(T-T_I)}{\alpha + \alpha'(T-T_I)}\right]} \tag{24}$$

In (24) $a = a(T(t))$ and $\alpha = \alpha(T(t))$.

The solution for the optimal path $T(t)$ can be generated by means of an iterative scheme, as noted in reference [1]. However for the AS/3502 and T300/5208 graphite/epoxy systems the data listed in Tables 1 and 2 indicate that (23) and (24) differ by at most 2 percent from expressions (10) and (11) which assume constant α. Consequently, we shall retain (12) as our solution for the optimal path even for the variable α case at hand. We note however, that although the optimal temperature paths remain unaffected by the variation of α, the thermal stresses do differ significantly.

In the particular circumstance that

$$\alpha(T) = \alpha_0 + \alpha_1 T \tag{25}$$

we obtain

$$\sigma(t) = E_0\left\{\int_0^t Q(\tau,t)\left[\alpha_0 + \alpha_1(2T(\tau) - T_I)\right]d\tau + K\alpha(T_0)(T_I - T_0)\,Q(0,t)\right\} \tag{26}$$

with the same $Q(\tau,t)$ as in (14). The stress immediately following the initial discontinuity is

$$\sigma(0^+) = E_0 t_0^{-q}\alpha(T_0)(T_I - T_0) \tag{27}$$

After the final discontinuity we have

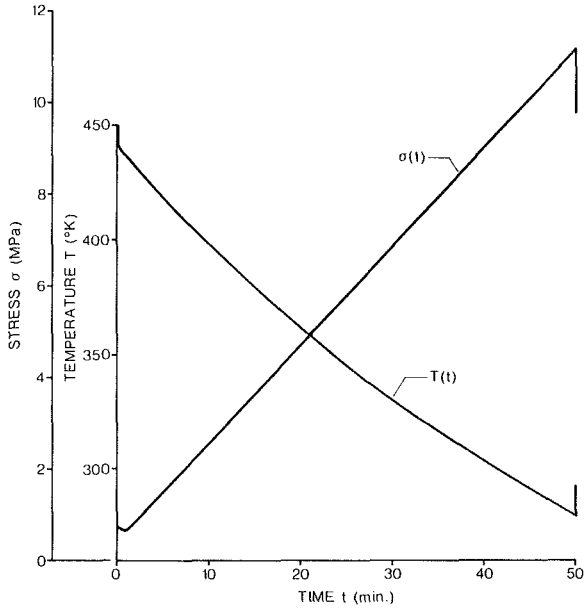


Fig. 1 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for epoxy/aluminum joints and cooling time t_f = 50 min

$$\sigma(t_f^+) = \sigma(t_f^-) + E_0 t_0^{-q}\left[\alpha(T_f)(T_f - T_I) - \alpha(T_R)(T_R - T_I)\right] \tag{28}$$

Computations

Calculations were performed for two graphite/epoxy materials and for an epoxy/aluminum joint. The data [3, 6] was reduced to conform with expressions (3), (7), (8), and (25) as shown in Tables 1 and 2.

In our computations we employed T_I = 450°K, T_R = 293°K, and considered cooling times t_f = 10, 50, 100, 500, and 1000 minutes.

The integrations (14) and (26) were performed numerically. For this purpose the time interval t_f was divided into n equal subintervals $\Delta t = t_f/n$, resulting in intermediate times $t_i = i\,\Delta t$ and intermediate temperatures T_i evaluated by means of (12). It was found that sufficient accuracy was obtained with $n = 100$. In order to assess the significance of the variation of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

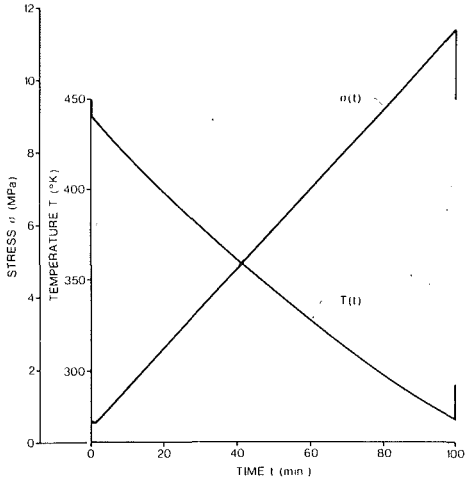


Fig. 2 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for epoxy/aluminum joints and cooling time t_f = 100 min

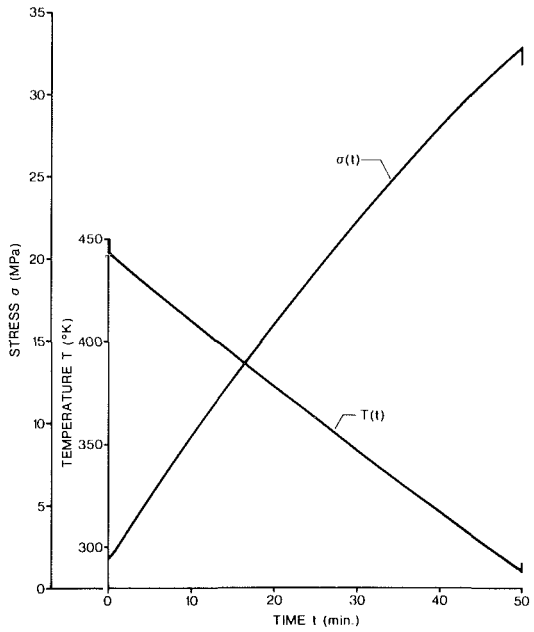


Fig. 3 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for T300/5208 symmetric, balanced 0/90 deg laminates and cooling time t_f = 50 min

α with T, we computed (15) with $\alpha_{av} = (\alpha(T_I) + \alpha(T_R))/2$. The results are shown in Figs. 1-9. In all those figures temperature T is in °K, stress is in MPa, and time is in minutes.

Figures 1 and 2 exhibit the optimal cooling paths $T(t)$, with the accompanying residual stress, generated during the cool-down process of typical aluminum/epoxy joints for cooling times t_f = 50 and 100 minutes, respectively. Note the jump discontinuities at $t=0$ and $t=t_f$. Analogous results are shown in Figs. 3 and 4 for the 0/90 deg T300/5208 laminate and in Figs. 5 and 6 for the 0/90 deg AS/3502 laminate.

In all the figures the optimal temperature paths were determined by (12). We note that those paths "undershoot" the level of room temperature (293°K), requiring an upward jump at $t = t_f$. The stresses in Figs. 1 and 2 were computed according to (15) and (16) employing constant α. On the other hand, the curves $\sigma(t)$ in Figs. 3-6 were obtained by employing (26)-(28) considering variable α in accordance with Table 2.

The variation of the residual thermal stress $\sigma(t_f^+)$ with the cooling time t_f is shown in Figs. 7-9. Also shown are the variations of the terminal temperature $T_f = T(t_f^-)$ with t_f. Note that a logarithmic scale is employed for t_f.

For comparison purposes we also exhibit by dashed-dotted

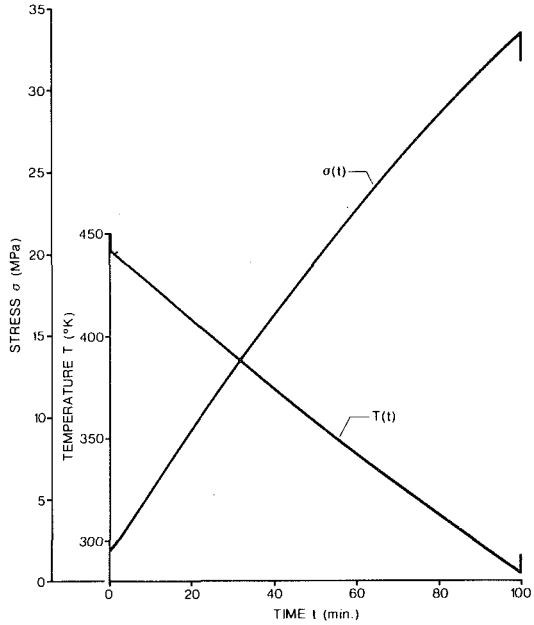


Fig. 4 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for T300/5208 symmetric, balanced 0/90 deg laminates and cooling time t_f = 100 min

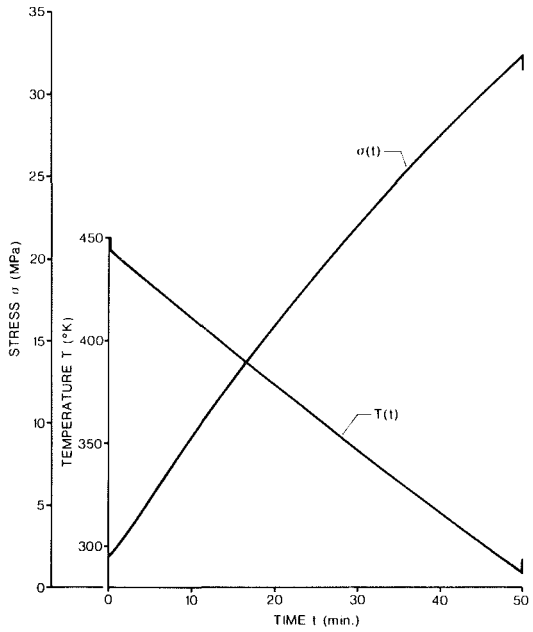


Fig. 5 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for AS/3502 symmetric, balanced 0/90 deg laminates and cooling time t_f = 50 min

lines in Figs. 8 and 9 the final residual stresses $\sigma(t_f^+)$ which result from considering a constant, average value of $\alpha_{av} = (\alpha(T_I) + \alpha(T_R))/2$ instead of the actual temperature-dependent $\alpha(T)$. Shown in dashed lines is the elastic value of $\sigma(t_f^+)$, which is computed with $\alpha = \alpha_{av}$.

We observe that by taking appropriate account of viscoelastic effects it is possible to attain reductions of about 25 percent in the residual stresses relative to the linear elastic predictions.

Remarks on Data Reduction

The "power law" representation of relaxation or creep data for polymeric resins or fiber-reinforced composites [3] in the form $E(t) = Et^{-q}$, can be surmised from the fact that the plots of log E versus log t are represented by straight lines [3,

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

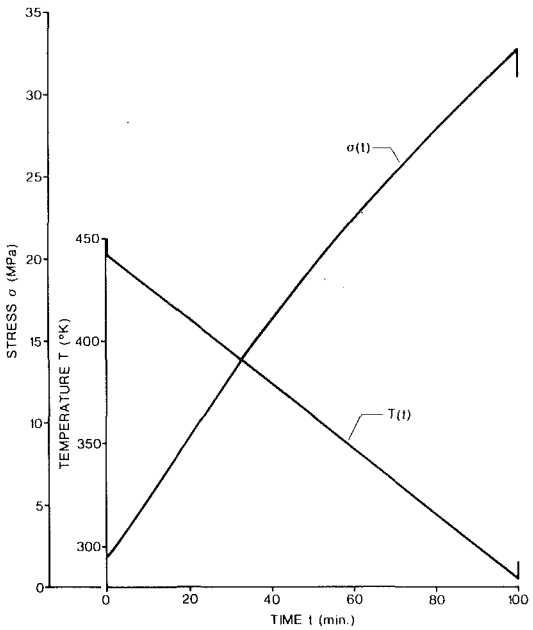


Fig. 6 The optimal cooling path $T(t)$ and the accompanying thermal stress $\sigma(t)$ versus time t for AS/3502 symmetric, balanced 0/90 deg laminates and cooling time t_f = 100 min

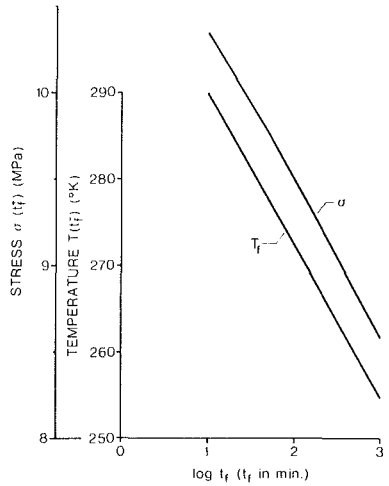


Fig. 7 The variation of $T(t_f^-)$ and the residual thermal stress $\sigma(t_f^+)$ versus log t_f for epoxy aluminum joints

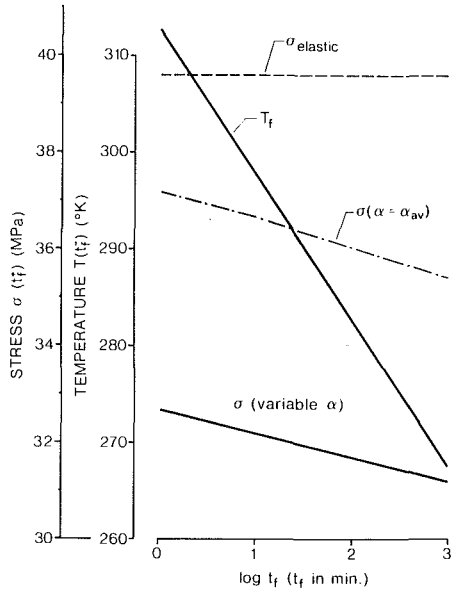


Fig. 8 The variation of $T(t_f^-)$ and the residual thermal stress $\sigma(t_f^+)$ versus log t_f for the T300/5208 0/90 deg laminates. (The dashed-dotted line corresponds to $\sigma(t_f^+)$ evaluated with $\alpha = \alpha_{av}$.)

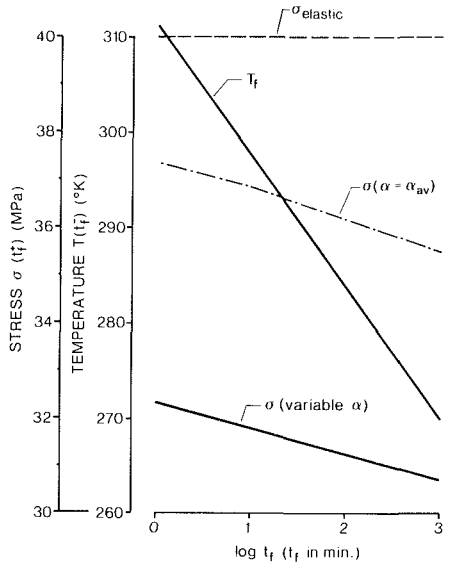


Fig. 9 The variation of $T(t_f^-)$ and the residual thermal stress $\sigma(t_f^+)$ versus log t_f for the AS/3502 laminates. (The dashed-dotted line corresponds to $\sigma(t_f^+)$ evaluated with $\alpha = \alpha_{av}$.)

6]. The slopes of those lines determine the power q. Obviously, these plots cannot be extrapolated to time $t = 0$, lest $E(0) \to \infty$. Furthermore, creep or relaxation data are unreliable at very short times because of the dynamic effects that are technically unavoidable at the early stages of experiments. Finally, it may also be noted that for $E(t) = ET^{-q}$ the optimization scheme (11) fails because it yields $\lim_{t \to \infty} t'(t) \to -\infty$

To overcome the aforementioned artifact, select a certain short time $t = t_s$ at which the data plots are smooth and beyond which $E(t) = Et^{-q}$. The data can then be extrapolated back to $t = 0$ employing Taylor's expansion and the result fitted into the form $E(t) = E_0(t+t_0)^{-q}$ of equation (6). It turns out that by letting $E_0 = E(1+q)$ and $t_0 = t_s$ we obtain representations in the form of (6) that match all data points indistinguishably from $E(t) = Et^{-q}$.

Acknowledgment

The authors wish to thank Mr. W. Eue of Texas A&M University for his help in obtaining the numerical results presented herein. This work was conducted under Contract F49620-78-C-0034 from the Air Force Office of Scientific Research (AFOSR) whose support is gratefully adknowledged.

References

1 Weitsman, Y., "Optimal Cool-Down in Linear Viscoelasticity," ASME Journal of Applied Mechanics, Vol. 47, 1980, pp. 35-39.

2 Gurtin, M. E., and Murphy, L. F., "Optimal Temperature Paths for Thermorheologically Simple Viscoelastic Materials," *Quart. App. Math.*, Vol. 38, No. 2, 1980, pp. 179-190.

3 Kibler, K. G., "Time-Dependent Environmental Behavior of Graphite/Epoxy Composites," Technical Report AFWAL-TR-80-4052, General Dynamics Corporation, 1980, Fort Worth, Texas.

4 Jones, R. M., *Mechanics of Composite Materials*, Scripta Book Co., 1975, pp. 37-41.

5 Tsai, S. W., and Hahn, H. T., *Introduction to Composite Materials*, Technomic, 1980, pp. 8-18.

6 Beckwith, S. W., "Viscoelastic Characterization of a Nonlinear Glass/Epoxy Composite Including the Effect of Damage," Ph.D. dissertation, 1974, Texas A&M University, College Station, Texas.

7 Morland, L. W., and Lee, E. H., "Stress Analysis for Linear Viscoelastic Materials With Temperature Variation," *Transactions of the Society of Rheology*, Vol. 4, 1960, pp. 233-260.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Y. A. Bahei-El-Din[1]
Lecturer,
Division of Structures,
Department of Civil Engineering,
Cairo University,
Giza, Egypt

G. J. Dvorak
Professor and Chairman,
Department of Civil Engineering,
University of Utah,
Salt Lake City, Utah 84112
Mem. ASME

Plasticity Analysis of Laminated Composite Plates

Elastic-plastic behavior of symmetric metal-matrix composite laminates is analyzed for the case of in-plane mechanical loading. The overall response of the laminate at each instant is derived from the elastic-plastic deformation of the individual fibrous layers, and from their mutual constraints. Constitutive equations of the laminated plates are presented in terms of initial yield conditions, hardening rules, and instantaneous compliances. Local stresses, hardening parameters, and strains are found in each lamina and in the fiber and matrix phases within each lamina. Specific results are obtained with the continuum model of elastic-plastic fibrous composites [1] which has been recently developed by the authors. Comparisons of analytical results with experimental measurements are made for certain laminated plates.

1 Introduction

Laminated plates consist of several layers of unidirectionally reinforced, fibrous composite laminae which have different in-plane orientations and are bonded together in a certain stacking sequence. When the matrix that contains the fibers is made of a soft metal, such as aluminum or magnesium, the laminated plate may experience plastic deformation in the course of loading. The reinforcing fibers usually have high stiffness and strength, and are elastic until failure. The failure strain of the fibers used in actual material systems varies from 0.005–0.015 [1]. In contrast, the as-fabricated metal matrices start to yield plastically at strains equal to 0.001 or so. Therefore, plastic straining in metal-matrix laminates may start at relatively low applied-load levels, and it may affect the mechanical response in a large part of the strength range of the composite material.

The elastic-plastic behavior of laminated plates is very different from that of unreinforced metallic plates under similar loading conditions. This is primarily caused by the interaction between the matrix and the fiber layers, which leads to constraint hardening effects that are absent in plasticity of homogeneous materials. To illustrate some of these effects, and to introduce the subject of the paper in a heuristic way, we present an example in which a (0/90) symmetric laminated plate made of a 6061-0 aluminum matrix and 30 volume percent of boron fiber is loaded by uniform in-plane normal stresses σ_{11}, σ_{22}. Figure 1 shows the loading and the reinforcement of the plate element. Using the results obtained in [1], and in the present paper, one can construct the initial yield surface of the laminated plate for the biaxial in-plane loading in the overall stress plane σ_{11}/Y, σ_{22}/Y. (Y is the matrix yield stress in simple tension.) The initial yield surface, which bounds the region of purely elastic material behavior, is the inner envelope of the local yield surfaces of individual laminae. These are plotted as two intersecting ellipses in the overall stress plane. When the composite plate is

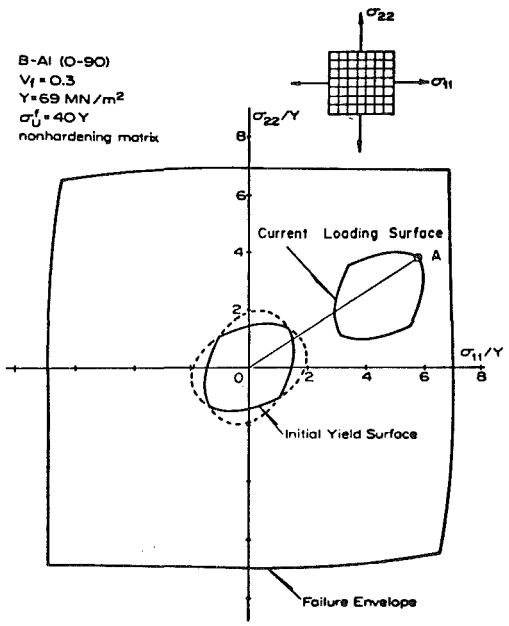


Fig. 1 The elastic-plastic deformation range of a biaxially loaded crossply *B-Al* plate

[1] Formerly: Visiting Assistant Professor, Department of Civil Engineering, Duke University, Durham, N.C.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November, 1981; final revision, March 1982. Paper No. 82-WA/APM-11.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

loaded beyond the elastic range, e.g., along the path OA in Fig. 1, each of the two local yield surfaces translates according to its own hardening rule, and in response to the part of the overall stress that is carried by the particular lamina. The current loading surface, which generally has a different shape than the initial yield surface, again bounds the region of elastic response. It also contains the current loading point A, which is found at a corner of the loading surface if both laminae yield simultaneously at A. It is seen that due to constraint hardening the loading surface has experienced a significant translation. The origin is now outside the loading surface, so that complete unloading would cause additional plastic straining. This is indeed observed in experiments with laminated composite plates which are discussed in the sequel (c.f. Fig. 8). Finally, Fig. 1 shows an approximate failure envelope, constructed for a family of proportional loading directions, where the overall stresses would cause local fiber failure in the plastically deforming plate. The tensile strength of the fiber was used in this calculation; the compression branches may underestimate actual behavior. However, it is clear that the strength range is considerably larger than the elastic range, and that loading into the plastic range is necessary for utilization of the benefits offered by the fiber reinforcement.

This paper presents a derivation of constitutive equations that describe the elastic-plastic behavior of laminated plates in the case of in-plane mechanical loading. We shall limit our attention to plates that are symmetric about their midplane, contain no imperfections or initial stresses, and are not affected by free edge effects. Under such circumstances, each lamina, as well as the laminated plate, will remain in the state of plane stress at each loading step. Analysis of more complex geometries and loading regimes is beyond the scope of the paper, but it can be accomplished with the finite element procedure described in [2]. The constitutive equations will be derived from those of the layers, with regard to their in-plane orientations, stress equilibrium, and mutual constraints. Specific forms will be developed with the model of elastic-plastic behavior of unidirectional lamina described in [1–4]. The results will be applied to selected laminates, and compared with experimentally observed behavior.

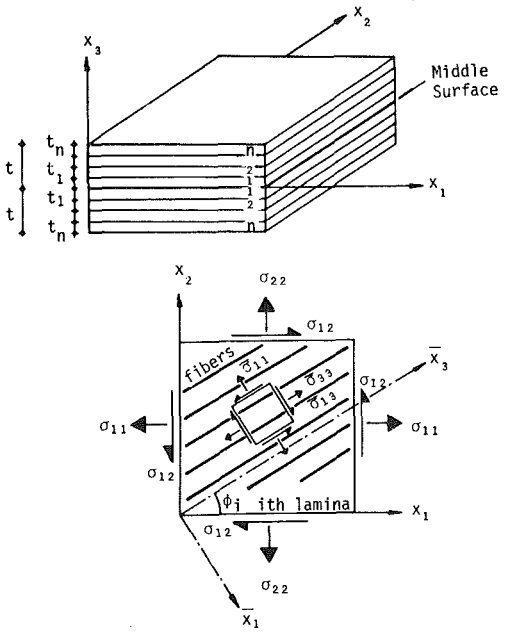


Fig. 2 Definition of geometry and coordinate systems of a composite laminate under in-plane stresses

2 Notation

We shall use a notation that is formally similar to that in [1], and which follows an established procedure in the theory of heterogeneous media introduced by Hill [5]. 3×3 matrices will be denoted by uppercase, lightface, Latin italicized letters, e.g., L, M; 3×1 matrices by boldface, lowercase, Greek letters, such as $\boldsymbol{\sigma}$, and $\boldsymbol{\epsilon}$; and scalars by lightface, lowercase italicized Latin, and Greek letters.

The overall in-plane properties of laminated plates will be derived from those of the laminae, which, in turn depend on the fiber and matrix properties within each lamina. To this end we introduce two coordinate systems, Fig. 2. The overall coordinates x_j ($j=1, 2, 3$) are associated with the laminated plate. The local coordinates $\bar{x}_k$ ($k=1, 2, 3$) are the material axes of the lamina, such that $\bar{x}_3$ coincides with the local fiber direction, and $\bar{x}_1$, $\bar{x}_2$ are in the transverse plane. These local coordinates are identical with the x_i system used in [1]. The x_1x_2-plane is the midplane of the laminated plate, the $\bar{x}_1\bar{x}_3$-planes are midplanes of the laminae, and are parallel to the x_1x_2-plane. The angle ϕ_i is contained by the fiber axis $\bar{x}_3$ in each lamina (i), and the x_1 axis of the plate.

The behavior of each lamina (i) will be described both in the local coordinates $\bar{x}_k$, and in the overall coordinates x_j. The description is analogous to that used in [1].

In the local system $\bar{x}_k$:

$$d\bar{\boldsymbol{\sigma}}_i=\bar{L}_i d\bar{\boldsymbol{\epsilon}}_i,\quad d\bar{\boldsymbol{\epsilon}}_i=\bar{M}_i d\bar{\boldsymbol{\sigma}}_i, \tag{1}$$

where

$$d\bar{\boldsymbol{\sigma}}_i=[d\bar{\sigma}_{11}^{(i)}\, d\bar{\sigma}_{33}^{(i)}\, d\bar{\sigma}_{13}^{(i)}]^T,\quad d\bar{\boldsymbol{\epsilon}}_i=[d\bar{\epsilon}_{11}^{(i)}\, d\bar{\epsilon}_{33}^{(i)}\, 2d\bar{\epsilon}_{13}^{(i)}]^T \tag{2}$$

are the local stress and strain-increment vectors in lamina (i), and $\bar{L}_i$, $\bar{M}_i$ are the instantaneous moduli and compliances, such that $\bar{M}_i=\bar{L}_i^{-1}$ if the inverse exists. The matrices $\bar{L}_i$ and $\bar{M}_i$ are symmetric for elastic and stable elastic-plastic materials [6, 7]. Equation (1) is identical with (6) in [1], except for the top bars on $\bar{L}_i$, $\bar{M}_i$, which have now been added. Also, only those coefficients of $\bar{M}_i$ derived in [1] which correspond to the plane stress state (2) need be considered. In particular,

$$\bar{M}_i=\begin{bmatrix}\bar{M}_{11} & \bar{M}_{13} & \bar{M}_{15}\\ & \bar{M}_{33} & \bar{M}_{35}\\ \text{Sym.} & & \bar{M}_{55}\end{bmatrix} \tag{3}$$

where the $\bar{M}_{ij}$ are given by (13) and (50) in [1], or by (A-6) in the Appendix.

In the overall system x_j:

$$d\boldsymbol{\sigma}_i=L_i d\boldsymbol{\epsilon}_i,\quad d\boldsymbol{\epsilon}_i=M_i d\boldsymbol{\sigma}_i, \tag{4}$$

where

$$d\boldsymbol{\sigma}_i=[d\sigma_{11}^{(i)}\, d\sigma_{22}^{(i)}\, d\sigma_{12}^{(i)}]^T,\quad d\boldsymbol{\epsilon}_i=[d\epsilon_{11}^{(i)}\, d\epsilon_{22}^{(i)}\, 2d\epsilon_{12}^{(i)}]^T \tag{5}$$

are the local stress and strain-increment vectors in the overall coordinates x_j.

From the geometry of the coordinate systems x_j and $\bar{x}_k$, Fig. 2, one can derive the following transformation relations:

$$d\bar{\boldsymbol{\sigma}}_i=T_i d\boldsymbol{\sigma}_i,\quad d\bar{\boldsymbol{\epsilon}}_i=N_i d\boldsymbol{\epsilon}_i \tag{6}$$

where

$$T_i^T=N_i^{-1}=\begin{bmatrix}\sin^2\phi_i & \cos^2\phi_i & \tfrac{1}{2}\sin2\phi_i\\ \cos^2\phi_i & \sin^2\phi_i & -\tfrac{1}{2}\sin2\phi_i\\ -\sin2\phi_i & \sin2\phi_i & -\cos2\phi_i\end{bmatrix}, \tag{7}$$

and ϕ_i is the angle contained by the axis x_1 and the fiber direction $\bar{x}_3$ in lamina (i). It will become obvious that the particular choice of the coordinate systems in Fig. 2, and consequently the forms of (2), (3), (5), and (7) have no effect on the generality of the results that follow.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The transformation relations for the instantaneous stiffnesses and compliances are:

$$L_I = N_i^T \bar{L}_i N_i \quad M_i = T_i^T \bar{M}_i T_i \tag{8}$$

Finally, for the laminated composite plate, in the overall system x_j, the constitutive equations will be written as:

$$d\sigma = L\, d\epsilon, \quad d\epsilon = M\, d\sigma, \tag{9}$$

where the vectors

$$d\sigma = [d\sigma_{11} d\sigma_{22} d\sigma_{12}]^T, \quad d\epsilon = [d\epsilon_{11} d\epsilon_{22} 2d\epsilon_{12}]^T, \tag{10}$$

are the macroscopic stress and strain increments applied to the laminated plate, and L, M are the instantaneous stiffnesses and compliances; $M = L^{-1}$ if the inverse exists. The evaluation of the unknowns L and M is our principal goal.

Throughout the paper we shall assume that the overall stresses and strains applied to the plate are uniform.

3 Governing Equations

Consider a laminated plate consisting of $2n$ layers. The layers are assumed to be placed symmetrically with respect to the midplane x_1x_2 of the plate, Fig. 2. The thickness of the ith lamina is denoted as t_i, $i = 1,2,...n$; the total thickness of the laminated plate is equal to $2t$. Then, the volume fraction of the ith layer is:

$$c_i = t_i/t, \quad \sum_{i=1}^{n} c_i = 1 \tag{11}$$

Since the layers are perfectly bonded together, the following constraints exist between increments (5) and (10):

$$d\epsilon = d\epsilon_1 = d\epsilon_2 = \ldots = d\epsilon_n \tag{12}$$

The local stress increments $d\sigma_i$ in the layers must be in equilibrium with the overall stress increment $d\sigma$:

$$d\sigma = c_1 d\sigma_1 + c_2 d\sigma_2 + \ldots + c_n d\sigma_n. \tag{13}$$

Also, we define stress distribution factors H_i for individual layers:

$$d\sigma_i = H_i d\sigma, \tag{14}$$

with the property following from (13):

$$c_1 H_1 + c_2 H_2 + \ldots + c_n H_n = I, \tag{15}$$

where I is a 3×3 identity matrix.

These relations permit the derivation of overall instantaneous properties of the laminated plate in terms of the local instantaneous moduli and compliances. The results are:

$$L = c_1 L_1 + c_2 L_2 + \ldots + c_n L_n \tag{16}$$

$$M = M_1 H_1 = M_2 H_2 = \ldots = M_n H_n; \tag{17}$$

and

$$H_i = L_i L^{-1} = M_i^{-1} M. \tag{18}$$

Since $d\epsilon$ and $d\sigma$ are uniform, it follows from (12), (14), and (6) that $d\epsilon_i$, $d\sigma_i$, and $d\bar{\epsilon}_i$, $d\bar{\sigma}_i$ are also uniform.

To relate the lamina stresses to the stress averages σ_r $(r = f,m)$ in the fiber and matrix constituents, we write

$$d\sigma_r = B_r d\bar{\sigma}, \tag{19}$$

where B_r is the phase (r) stress-concentration factor, as in [1]. In the elastic range $B_r = B_{re}$. Then, the local stress increments in phase (r) of lamina (i) in terms of the overall stress increments applied to the laminated plate are:

$$d\sigma_r^{(i)} = R_r^{(i)} d\sigma, \quad (r = f,m); \tag{20}$$

and, from (6), (14):

$$R_r^{(i)} = B_r^{(i)} T_i H_i. \tag{21}$$

If $B_r^{(i)} = B_{re}^{(i)}$, then $R_r = R_{re}^{(i)}$, $r = f,m$.

The stress-concentration factors B_r, $R_r^{(i)}$ satisfy the connections:

$$c_f B_f + c_m B_m = I, \quad c_f^{(i)} R_f^{(i)} + c_m^{(i)} R_m^{(i)} = T_i H_i \tag{22}$$

c_f and c_m denote the phase volume fractions, $c_f + c_m = 1$.

4 Elastic-Plastic Behavior

4.1 Initial Yielding. We assume that the initial yield condition for each lamina is known, and that it can be written in the form

$$f_i(\sigma_i) = 0, \tag{23}$$

where σ_i are local stresses in the ith lamina in the overall coordinate system x_j. Each of the local yield surfaces can be expressed in terms of the overall stresses σ:

$$f_i(H_i^e \sigma) = 0, \tag{24}$$

where H_i^e is the elastic stress-distribution factor (18).

The overall yield surface for in-plane loading of the laminated plate in the stress space σ is the internal envelope of the lamina yield surfaces (24). If the surfaces (23) are convex, then (24) are also convex, and they may intersect at a finite number of corners. Some, and possibly all, of these corners may become part of the overall yield surface.

A specific form of (23) may be obtained with the procedure outlined in references [1] and [3]. There, a continuum model of a unidirectional fibrous composite was developed from the assumption that the elastic fibers, which are embedded in the elastic-plastic matrix, have vanishing diameters and that they

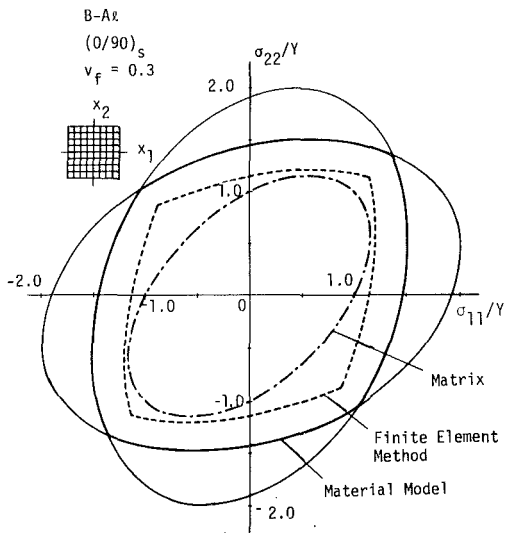


Fig. 3 Initial yield surface of a crossply *B-Al* plate

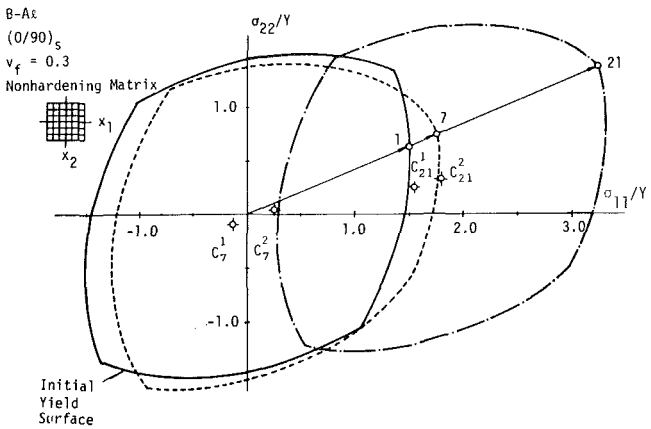


Fig. 4 Loading surfaces of a crossply *B-Al* plate

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

occupy a finite volume fraction of the composite. The model retains the essential axial constraint between the phases. The local stress and strain fields in the phases are uniform. The lamina yield condition from this model with a Mises-type matrix is given by equation (22) in [1]; it can be written as:

$$f(\bar{\sigma}) = \bar{\sigma}^T \bar{K} \bar{\sigma} - Y^2 = 0. \tag{25}$$

Here, $\bar{\sigma}$ are the lamina stresses in coordinates $\bar{x}_k$, the matrix $\bar{K}$ depends on the elastic moduli of the fiber and matrix, and on the phase volume fractions; Y is the matrix yield stress in simple tension. An explicit form of (25) is given by (A-1) in the Appendix.

The lamina yield surface in the overall coordinates x_j is, from (6), (23), and (25):

$$f_i(\sigma_i) = [\sigma^T T^T \bar{K} T \sigma]_{(i)} - Y_i^2 = 0 \tag{26}$$

In terms of overall stresses σ:

$$f_i(\sigma) = \sigma^T [(H^e)^T T^T K T H^e]_{(i)} \sigma - Y_i^2 = 0 \tag{27}$$

or, as in (25):

$$f_i(\sigma) = \sigma^T K_i \sigma - Y_i^2 = 0, \tag{28}$$

with

$$K_i = (H_i^e)^T T_i^T \bar{K}_i T_i H_i^e. \tag{29}$$

The internal envelope of the $f_i(\sigma)$ family of lamina yield surfaces is the overall initial yield surface of the composite plate. Here, and in the sequel, we denote by Y_i the matrix tensile yield stress in the lamina (i), even though in most applications there is $Y_i = Y$ for all i.

Figure 3 shows an example of the overall initial yield surface for the crossply boron-aluminum plate which was introduced in Fig. 1. Also presented in Fig. 3 is the Mises yield surface of the unreinforced matrix, and the initial yield surface obtained for the plate in reference [8] from a finite element analysis of a laminate model which accounted for the effects of fiber crossover at interfaces between the layers, and thus provided an estimate of the onset of microscale yielding in the composite. In constrast, the present analysis, and the material model [1-3], estimate the onset of macroscopic yielding in the composite. However, the two yield surfaces of the laminated plate are of similar size, which indicates that under plane-stress load conditions, plastic yielding in the matrix may develop very rapidly, within a small stress interval, after the localized yielding sets in. Furthermore, it is seen that the initial yield surface of the unreinforced matrix is not appreciably smaller than that obtained for the laminated plate. That indicates that the fiber reinforcement does not cause a significant expansion of the elastic domain of the matrix.

4.2 Subsequent Yield Surfaces. When the laminated plate is loaded by applied stresses σ in such a way that $f_i(\sigma) < 0$ in (27), the plate remains elastic and its overall properties can be evaluated directly from (16), with the lamina stiffnesses L_i taken from (13) in [1], or as an inverse of (3). If the loading point reaches one of the branches (i) of the internal envelope of (27) or (28), then $f_i(\sigma) = 0$, and the lamina (i) will experience the onset of plastic yielding. Other layers of the laminate will remain elastic, unless the loading point σ coincides with a corner of the yield surface where two or more branches intersect. If that is the case, the onset of yielding will take place in each lamina that has a branch of the overall yield surface containing the loading point.

When loading continues beyond initial yielding into the plastic range, the laminate yield surface must deform to contain the loading point. In analogy with the terminology used in plasticity of homogeneous media, the change in the shape and position of the yield surface in the process of plastic loading will be referred to as hardening. In plastic deformation of composite media, there are, in general, two sources

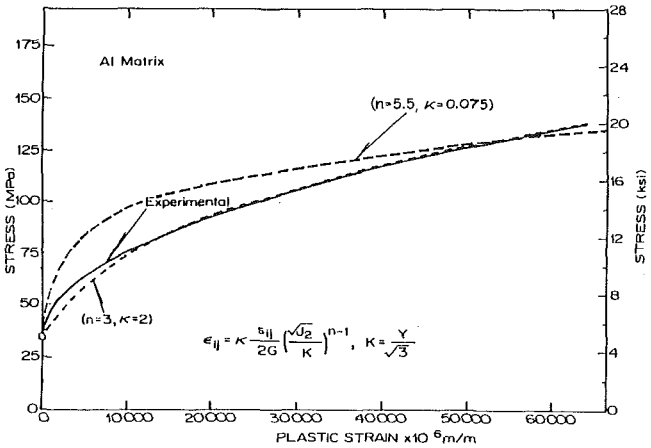


Fig. 5 Ramberg-Osgood approximation of matrix stress-strain curve

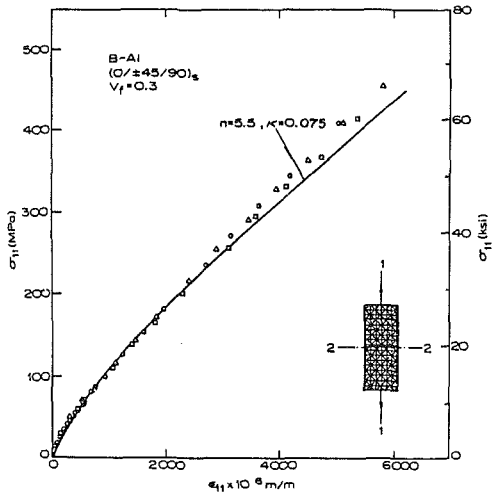


Fig. 6 Comparison of calculated and measured stress-strain curves for a *B-Al* plate

of hardening. One associated with hardening of each of the constituent phases, and another that is caused by the mutual constraint between the bonded phases. These are usually called phase hardening and constraint hardening, respectively.

In the case of laminated plates, the term constraint hardening will be reserved for the effects that arise from the constraint between the composite layers. The layers will be regarded as phases, with certain phase-hardening rules. Of course, the hardening of each fibrous lamina is also composed from phase (matrix) and constraint (fiber/matrix) components [1]. The constraints at the different levels usually have the dominant influence on overall behavior. Hardening of the matrix phase is relatively unimportant. For aluminum or magnesium matrices it can be adequately described by the Mises yield function with the Prager-Ziegler kinematic hardening rule, as in [1].

To obtain a specific hardening rule for the laminated plate we shall restrict our attention to the class of composite models that lead to kinematic phase hardening. We recall from [1] that kinematic hardening must always be present in fiber-reinforced materials, as a result of local constraint hardening of the lamina. Other types of lamina hardening are possible, but not necessary, and will be excluded in the present analysis. As a special case one may admit layers that are not reinforced, but which harden kinematically. Such layers may be encountered in sandwich plates which consist of both reinforced and unreinforced sheets.

Specifically, we shall consider only the following form of phase hardening in laminae:

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$f_i(\sigma_i - \alpha_i) = 0. \tag{30}$$

In analogy with (24), the overall yield surface of the laminated plate, consisting of laminae that have the property (30), will be

$$f_i(\sigma - \hat{\alpha}_i) = 0 \tag{31}$$

The vector $\hat{\alpha}_i$ denotes the translation of the yield surface of the ith lamina in the overall stress space σ.

To determine $\hat{\alpha}_i$, we assume that the stress vector σ satisfies (24), or (31), and that a stress cycle $\pm d\sigma$ has been applied such that $(\partial f_i/\partial\sigma)\cdot d\sigma > 0$ in the $+d\sigma$ part of the cycle. The $-d\sigma$ part causes elastic unloading. At the end of the cycle, each lamina (i) will contain a residual stress increment

$$d\sigma_i^p = (H_i - H_i^e)\, d\sigma. \tag{32}$$

Also, for each lamina (i) which has experienced plastic deformation during the loading cycle, there will be a local translation $d\alpha_i$ of the local yield surface (30).

The residual stresses $d\sigma_i^p$ and the local translations $d\alpha_i$ are equivalent to a local stress change that reduces the magnitude of the overall stress σ at yield. This change is accommodated by a translation of the surfaces (27) by the amount

$$d\hat{\alpha}_i = -(H_i^e)^{-1}(d\sigma_i^p - d\alpha_i). \tag{33}$$

Using (32), one obtains

$$d\hat{\alpha}_i = (H_i^e)^{-1} d\alpha_i - [(H_i^e)^{-1} H_i - I] d\sigma, \tag{34}$$

where I is a 3×3 identity matrix. Since the phase-hardening rules (26) are usually given in the local coordinates x_k, in the form

$$f_i(\bar{\sigma}_i - \bar{\alpha}_i) = 0, \tag{35}$$

it is useful to rewrite (34), with the help of (6) and (7), as

$$d\hat{\alpha}_i = (T_i H_i^e)^{-1} d\bar{\alpha}_i - [(H_i^e)^{-1} H_i - I] d\sigma. \tag{36}$$

Here, the first term represents the contribution of phase hardening, and the second the effect of constraint hardening on $d\alpha_i$. The latter is present even in laminae that have not yielded and in which $d\bar{\alpha}_i = 0$.

The translation vectors $d\hat{\alpha}_i$ in (31) can be evaluated by integration of (36) along the loading path. Since the H_i depend on the instantaneous overall moduli L, or compliances M, and H_i^e on the elastic overall properties, the local yield surfaces (31) will not translate independently of each other. Rather, they will always assume a clustered configuration with an internal envelope that will be the current yield surface of the laminate, and which will contain the current loading point.

4.3 Overall Strains. The overall compliance matrix M (17), which is required in evaluation of the overall strain increment $d\epsilon$ in (9), must be found in conjunction with the solution of a system of equations that describe the response of each lamina, of the fiber and matrix in each lamina, and also of the laminate itself at each step of plastic loading. These equations, some of which were derived in [1], will now be rewritten in a form more suitable for their numerical solution.

Suppose that the composite plate is loaded by a current overall stress σ such that (31) is satisfied in at least one lamina (i). Let a stress increment

$$d\sigma = d\sigma^* \tag{37}$$

be applied such that

$$(\partial f_i/\partial\sigma)\cdot d\sigma^* > 0 \tag{38}$$

and (31) are satisfied simultaneously in a certain number of laminae. We shall refer to these as yielding laminae, whereas those that do not deform plastically ($f_i < 0$ in (31)) in step (37) will be called elastic laminae.

To determine if (31) and (38) are satisfied, one must examine specific forms of (24) and (23), e.g., equations (25)–(27) when one adopts the composite model developed in [1]. In any event, since the fibers remain elastic, the lamina yield condition must be evaluated from that of the matrix. A general form, which includes (25) as a special case, is given by equation (21) in [1]:

$$f(B_{me}\bar{\sigma}) = 0, \tag{39}$$

where B_{me} is the elastic stress-concentration factor for the matrix phase (19).

From (20) and (37) one obtains the local stress increments as

$$d\sigma_m^{(i)} = R_m^{(i)} d\sigma^*, \quad d\sigma_f^{(i)} = R_f^{(i)} d\sigma^*; \tag{40}$$

and the lamina yield condition (23) as

$$f_i(R_{me}^{(i)}\sigma) = 0 \tag{41}$$

Evaluation of $d\hat{\alpha}_i$ in (34) must also be related to matrix hardening within each lamina. In accordance with the limitations adopted in (30) we can admit only kinematically hardening matrices in the present analysis. We use a Mises-type matrix which obeys Ziegler's kinematic hardening rule (equation (29) in [1]):

$$d\alpha_m = d\mu_m(\sigma_m - \alpha_m) \tag{42}$$

with $d\mu_m$ obtained in equation (34) in [1] as:

$$d\mu_m = (1/Y^2)(\sigma_m - \alpha_m)^T C d\sigma_m. \tag{43}$$

The 3×3 matrix C comes from the Mises yield condition, it contains only numerical constants: $C_{11} = C_{22} = 1$, $C_{33} = 3$, $C_{12} = -1/2, C_{13} = C_{23} = 0$; $C = C^T$.

Hence, for each yielding lamina (i):

$$d\alpha_m^{(i)} = [(1/Y^2)(\sigma_m^{(i)} - \alpha_m^{(i)})^T C\, R_m^{(i)} d\sigma^*](\sigma_m^{(i)} - \alpha_m^{(i)}) \tag{44}$$

Thus (34) becomes:

$$d\hat{\alpha}_i = (R_{me}^{(i)})^{-1} d\alpha_m^{(i)} - [(R_{me}^{(i)})^{-1} R_m^{(i)} - I] d\sigma^* \tag{45}$$

The evaluation of the instantaneous overall compliance M, and stiffness L of the laminated plate requires construction of a solution of the system of equations (40), (44), (45), (16), and (9). The unknowns $d\sigma_m^{(i)}, d\sigma_f^{(i)}, d\alpha_m^{(i)}, d\hat{\alpha}_i$, and $d\epsilon$ are functions of current values of $\sigma_m^{(i)}$ and $\alpha_m^{(i)}$; this is seen from (44), (45), and $(A-3)$–$(A-6)$. When the overall applied loading is proportional, e.g., in the form $d\sigma^* = \beta d\sigma_{11}$, the system can be solved with the help of the second Runge-Kutta formula, as shown in [4]. More complex loading conditions may require the use of a finite element procedure [2].

Since the plasticity analysis of the unidirectional lamina is available only in terms of compliances $\bar{M}_{ij}$(3), it is often necessary to calculate the stiffnesses L_i in (16)–(18) as $L_i = M_i^{-1}$. The compliance M_i becomes unbounded for certain loading directions in laminae with nonhardening matrices ($H_m = 0$ in (A-4)), L_i is then singular, and it cannot be evaluated from M_i. However, the analyis can be applied to composites with nonhardening matrices, providing that the loading directions in which certain $\bar{M}_{ij} \to \infty$ are avoided. In the plane-stress case the only proscribed loading direction is the proportional path $\bar{\sigma}_{11} = \bar{\sigma}_{33} = 0$, $\bar{\sigma}_{13} \neq 0$.

5. Applications

As an example of laminate hardening we consider the problem of biaxial elastic-plastic deformation of a crossply *B-Al* plate which was discussed in connection with Fig. 3. The aluminum matrix is assumed to be nonhardening of the Mises type. Under proportional in-plane loading we trace the motion of the yield surface of the composite plate. An illustration of calculated results is shown in Fig. 4. There, we start with the initial yield surface in the overall stress indicated by the solid line, which has been transferred from Fig. 3. The

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

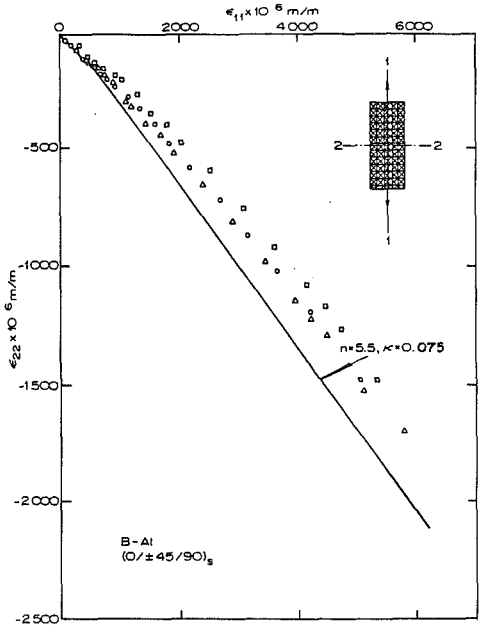


Fig. 7 Comparison of calculated and measured strains for a *B-Al* plate

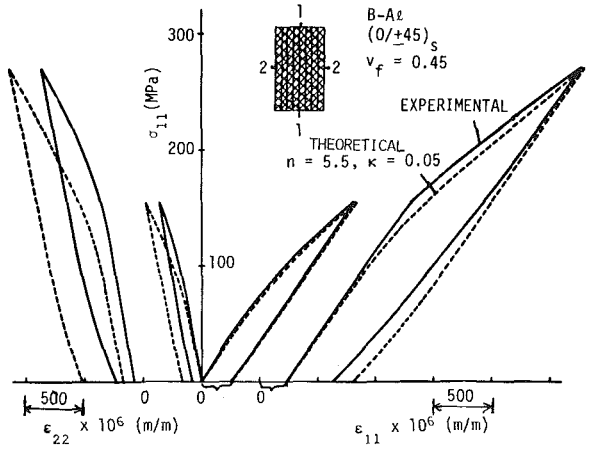


Fig. 8 Comparison of calculated and measured stress-strain curves for a $(0/\pm 45)_S$ *B-Al* plate subjected to two loading cycles

loading path is indicated by the straight line connecting points 1, 7, and 21. At loading point 1 the centers of the elliptic yield surfaces coincide with the origin. At this stress level, only the 90 deg lamina, with fibers perpendicular to the σ_{11} direction, has yielded. As the laminate is loaded gradually to loading Step 7, each of the two ellipses is translated in the overall stress space. Their respective centers are now denoted by C_7^1 and C_7^2. C^1 is the center of the yield surface of the 0 deg ply with fibers in the σ_{11} direction. Conversely, C^2 is the center of the yield surface of the 90 deg ply. At the loading Step 7, yielding starts in the 0 deg ply, in addition to the continued plastic deformation of the 90 deg ply. The laminate yield surface is obtained at this level of loading as the inner envelope of the translated laminae surfaces and is indicated by the dotted line in Fig. 4. The loading point 7 now coincides with a corner of the laminate yield surface. The surface assumes a new shape, different from the initial yield surface, due to the relative translation of the two elliptic branches.

Another position and shape of the overall yield surface is indicated at the loading point 21. The two centers of the laminae yield surfaces at this loading point are located at C_{21}^1 and C_{21}^2. Unloading of the laminate from point 21 along an arbitrary path to the origin would cause plastic deformation. Such inelastic unloading is commonly observed in experiments on laminated metal-matrix composite plates.

6 Comparison With Experiments

Calculations for laminated plates were made using the analysis outlined in the foregoing and compared to actual experimental measurements. The experiments were reported in [9].

In the experiments a group of laminated plates of different layup were tested in simple tension. The composite was made of a 6061 aluminum matrix reinforced by 43 volume percent of boron fibers. The stress-strain curve obtained experimentally for the aluminum material in the as-fabricated condition is shown in Fig. 5.

To define the properties of the aluminum matrix in the plastic range, the experimental curve in Fig. 5 was approximated by a Ramberg-Osgood type relation. However, the curve that fits best the behavior of unreinforced matrix, $n=3$, $\kappa=2$ in Fig. 5, provides a poor approximation of the subsequent experimental data. It was found that another material stress-strain law, with $n=5.5$, $\kappa=0.075$ in Fig. 5, gave better results. The reasons for this discrepancy are believed to be related to the effect of fibers on matrix hardening in the transverse plane, to in situ hardening characteristics of the matrix, and to other factors not accounted for in the material model used in the present calculations.

Using the two curves indicated in Fig. 5, calculations were made for several laminated plates. Figures 6 and 7 indicate how the calculated and experimental results compare in a laminated plate. The response in simple tension is shown in Fig. 6, and the comparison of axial and transverse strains is shown in Fig. 7.

Finally, Fig. 8 shows a comparison of experimental results [10] with finite element calculations [2] based on the present analysis, for the case of cyclic loading of a *B-Al* plate. Responses to each of the two cycles are drawn separately. The experiments confirm the theoretically predicted existence of plastic yielding after unloading, which is observed in the second cycle in Fig. 8. This behavior is in agreement with the concept of kinematic motion of the overall yield surface, illustrated in Fig. 4. The calculations predict quite well the laminate strains in the loading direction. However, they overestimate the transverse strains. This behavior may be traced to the choice of the material model of the lamina in [1], which neglects constraint between the fiber and matrix phases in the transverse plane.

7 Conclusions

The elastic-plastic response of composite laminates to in-plane loads has been determined from the local response of the individual laminae and their mutual constraints. The analysis presented here is applicable to large sheets of laminates containing no imperfections. Although interlaminar stresses that may develop locally at the free ends of laminates are not accounted for, comparison of analytical and experimental results shows reasonably close agreement. The comparison with experiments, however, has been obtained when matrix plastic properties different from those of an unreinforced material were used. The new properties are believed to be an approximation of the in situ properties of the matrix and have been obtained empirically.

The analysis can be used to estimate the response of binary or hybrid composites. Application of the analysis include prediction of failure loads and shakedown ranges for laminates [11].

Acknowledgments

This work has been supported by the U.S. Army Research Office.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

1 Dvorak, G. J., and Bahei-El-Din, Y. A., "Plasticity Analysis of Fibrous Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 327-335.

2 Bahei-El-Din, Y. A., Dvorak, G. J., and Utku, S., "Finite-Element Analysis of Elastic-Plastic Fibrous Composite Structures," *Computers and Structures*, Vol. 13, 1981, pp. 321-330.

3 Dvorak, G. J., and Bahei-El-Din, Y. A., "Plasticity of Composite Laminates," *Proceedings of Research Workshop on Mechanics of Composite Materials*, Dvorak, G. J., ed., Duke University, Oct. 1978, pp. 32-54.

4 Bahei-El-Din, Y. A., "Plastic Analysis of Metal Matrix Composite Laminates," Ph.D. Dissertation, Duke University, 1979.

5 Hill, R., "Elastic Properties of Reinforced Solids: Some Theoretical Principles," *Journal of the Mechanics and Physics of Solids*, Vol. 11, 1963, pp. 357-372.

6 Hill, R., "Continuum Micro-Mechanics of Elastoplastic Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 13, 1965, pp. 89-101.

7 Hill, R., "The Essential Structure of Constitutive Laws for Metal Composites and Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 15, 1967, pp. 79-95.

8 Rao, M. S. M., "Plasticity Theory of Fiber-Reinforced Metal Matrix Composites," Ph.D. Dissertation, Duke University, 1974.

9 Shukow, S. I., "Acoustic-Mechanical Characterization of Boron-Aluminum Composite Laminates," M. Sc. Thesis, The University of Delaware, 1977.

10 Sova, J. A., and Poe, C. C., Jr. "Tensile Stress-Strain Behavior of Boron-Aluminum Laminates," NASA-TP-1117, 1978.

11 Dvorak, G. J., and Johnson, W. S., "Fatigue of Metal Matrix Composites," *International Journal of Fracture*, Vol. 16, No. 16, 1980, pp. 585-607.

APPENDIX

Selected results obtained in [1] are reduced here to the case of plane stress, with the stresses given by (2). When referring to specific equations in [1] we follow the original equation number with [1], e.g., (35[1]).

The matrix K in (25) is obtained from (22[1]) as

$$K = \left[\begin{array}{c|c} (\mathbf{b}_m')^T \mathbf{c}\, \mathbf{b}_m' & 0 \\ \hline 0 & 3 \end{array}\right] \qquad (A\text{-}1)$$

where

$$\mathbf{b}_m' = \frac{1}{E_c}\begin{bmatrix} E_c & 0 \\ (1-c_m)a_m & E_m \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix},$$

$$E_c = c_f E_A^f + c_m E_m, \quad a_m = (\nu_m E_A^f - \nu_A^f E_m),$$

$$c_f + c_m = 1$$

Here, E_A^f, ν_A^f are axial Young's modulus and axial Poisson's ratio of a transversely isotropic fiber, E_m, ν_m are matrix elastic constants, and c_f, c_m are phase volume fractions in a lamina.

The elastic stress concentration factors B_{re} in (19) are, from (14[1]):

$$B_{re} = \left[\begin{array}{c|c} \mathbf{b}_r' & 0 \\ \hline 0 & 1 \end{array}\right], (r=f, m), \qquad (A-2)$$

with $\mathbf{b}_r'$ as in (A-1) when $r=m$. For $r=f$, B_{re} follows from (22).

In a yielding lamina, the instantaneous stress concentration factor of the matrix is, from (48[1])

$$B_m = \begin{bmatrix} 1 & 0 & 0 \\ B_{31}^m & B_{33}^m & B_{35}^m \\ 0 & 0 & 1 \end{bmatrix}, \qquad (A\text{-}3)$$

where

$$B_{31}^m = (c_f/h)\,(H_m a_m - E_A^f E_m \eta_1)$$
$$B_{33}^m = H_m E_m / h \qquad (A-4)$$
$$B_{35}^m = -(c_f/h)\,(E_A^f E_m \eta_5)$$

with $h = H_m E_c + c_f E_A^f E_m \eta_3$; H_m is the matrix-hardening parameter.

The coefficients η_1, η_3, and η_5 follow from (41[1]):

$$\eta_1 = q[2(\sigma_{11}^m - \alpha_{11}^m) - (\sigma_{33}^m - \alpha_{33}^m)]$$
$$\eta_3 = q[2(\sigma_{33}^m - \alpha_{33}^m) - (\sigma_{11}^m - \alpha_{11}^m)]$$
$$\eta_5 = 6q(\sigma_{13}^m - \alpha_{13}^m) \qquad (A\text{-}5)$$

and from (42[1]):

$$q = \frac{[-(\sigma_{11}^m - \alpha_{11}^m) + 2(\sigma_{33}^m - \alpha_{33}^m)]}{6[Y^2 + 3(\sigma_{13}^m - \alpha_{13}^m)^2]}.$$

The instantaneous compliances (3) for a yielding lamina are, from (50[1]):

$$\bar{M}_{11} = (c_m/E_m + c_f/E_T^f) - s\, a_m B_{31}^m + s\,\eta_1 (c_f a_m - E_c B_{31}^m)/(c_f \eta_3)$$
$$\bar{M}_{13} = -\nu_A^f/E_A^f - [c_m/(c_f E_A^f)] B_{31}^m$$
$$\bar{M}_{15} = s\,\eta_5 (c_f a_m - E_c B_{31}^m)/(c_f \eta_3) \qquad (A-6)$$
$$\bar{M}_{33} = (1 - c_m B_{33}^m)/(c_f E_A^f)$$
$$\bar{M}_{35} = s\,\eta_5 (E_m - E_c B_{33}^m)/(c_f \eta_3)$$
$$\bar{M}_{55} = (c_m/G_m + c_f/G_A^f) - s\,\eta_5 E_c B_{35}^m/(c_f \eta_3),$$

where $s = c_m/(E_m E_A^f)$, and E_T^f, G_A^f are the transverse Young's modulus, and axial shear modulus of the transversely isotropic fiber.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

G. Herrmann
Division of Applied Mechanics,
Stanford University,
Stanford, Calif. 94305
Fellow ASME

M. Hemami[1]
Isfahan, Iran

Plane-Strain Surface Waves in a Laminated Composite

Elastic wave propagation in plane strain in a periodically layered half space is considered, with the layering parallel to the bounding traction-free plane. Attention is focused on surface waves of the Rayleigh-type propagating parallel to the bounding plane. It is found that such waves are highly dispersive and that higher branches may be discontinuous.

Introduction

We are concerned here with a periodically layered half space which is formed by an infinite number of cells and bounded by a traction-free plane which is parallel to the layering, whose location with respect to the cell structure, however, is arbitrary. Each cell consists of two layers of different materials and different thicknesses, perfectly bonded together, which are, however, assumed to be homogeneous, isotropic, and linearly elastic.

In considering propagation of horizontally polarized shear waves (i.e., *SH* waves involving a state of antiplane strain) it has been recently shown [1-3] that there exist a new class of surface waves which, like Love waves, propagate along the free plane of the layered composite described in the foregoing. A method was presented for deriving these horizontally polarized shear-wave solutions from plane-wave solutions by applying Floquet's theorem to an infinite periodically layered medium [4].

It is the purpose of the present work to consider the propagation of compressional and vertically polarized shear waves (i.e., *P* and *SV* waves involving a state of plane strain). The work is based on results of recent analyses of wave solutions in plane strain obtained for an infinite layered medium, again by applying Floquet's theorem [5-7] and summarized in the next section.

Emphasis is placed on determining properties of Rayleigh-type surface waves propagating parallel to the bounding plane. For this purpose, an infinite periodically layered medium is considered first and the associated dispersion equation is briefly discussed. Next, the condition is imposed that some plane parallel to the layering be traction-free, which leads to the problem of reflection at a free boundary. Different types of waves in a half-space are discussed in the folowing section, and then narrowed down to a consideration of surface waves. The presentation concludes with an analysis of the properties of the lowest, and thus most interesting branch of the surface wave spectrum.

Periodically Layered Medium

The system considered consists of an infinite sequence of two alternating layers, each of which is linearly elastic, homogeneous, isotropic, and perfectly bonded to the adjoining layers. A unit cell is defined as the union of any two adjacent layers. The properties of two lamellae of a typical unit cell will be distinguished by using simple (unprimed) and primed notation: Lamé's elastic constants λ, μ; λ', μ', thicknesses $2h$, $2h'$ and densities ρ; ρ' (Fig. 1).

In the sequel, the numerical values assigned to these constants will be such that the phase velocities in the primed layer will be greater than in the unprimed layer. For this reason the primed layer will be called the fast layer, while the unprimed layer will be called the slow layer.

The most general two-dimensional motion in such a medium uncouples in motions in plane strain and in antiplane strain. Each of these has been discussed recently in [4-7], and the dispersion equations have been derived. In the following, steps leading to the derivation of the dispersion equation in plane strain are summarized first.

Let us consider a typical unit (*N*th) cell and write the equations of motion for each layer in its own local coor-

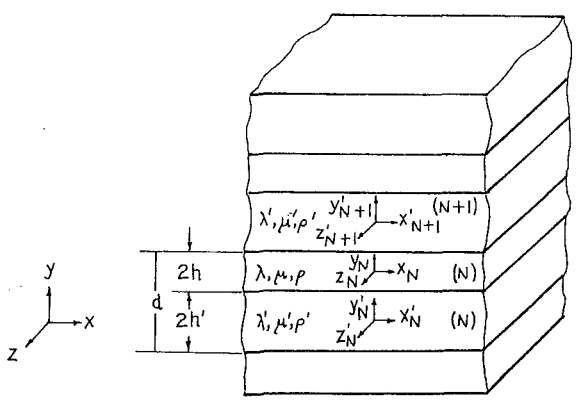


Fig. 1 *N*th unit cell

[1] Formerly, Research Assistant, Division of Applied Mechanics, Stanford University, Stanford, Calif. 94305.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion of this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May, 1981; final revision, May, 1982. Paper No. 82-WA/APM-21.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

dinates. For the unprimed Nth layer we have the displacements

$$u=u(x,y_N,t) \quad -h\le y_N\le h$$
$$v=v(x,y_N,t) \tag{1}$$

In terms of the displacement potential φ and ψ

$$u=\frac{\partial\varphi}{\partial x}+\frac{\partial\psi}{\partial y_N}$$
$$v=\frac{\partial\varphi}{\partial y_N}-\frac{\partial\psi}{\partial x} \tag{2}$$

The equations of motion are

$$(\lambda+2\mu)\left(\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y_N^2}\right)=\rho\ddot{\varphi}$$
$$\mu\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y_N^2}\right)=\rho\ddot{\psi} \tag{3}$$

and their solution may be written as

$$\varphi=\left(C_1\,e^{i\pi\beta y_N/2h}+C_2\,e^{-i\pi\beta y_N/2h}\right)$$
$$\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right] \quad P\text{ waves}$$
$$\psi=\left(C_3\,e^{i\pi\alpha y_N/2h}+C_4\,e^{-i\pi\alpha y_N/2h}\right)$$
$$\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right] \quad SV\text{ waves} \tag{4}$$

where

$$c_1=\left(\frac{\lambda+2\mu}{\rho}\right)^{1/2}$$
$$c_2=\left(\frac{\mu}{\rho}\right)^{1/2}$$
$$\Omega=\frac{2h}{\pi}\frac{\omega}{c_2}$$
$$\zeta=\frac{2h}{\pi}k_x$$
$$\alpha=\sqrt{\Omega^2-\zeta^2}=\frac{2h}{\pi}\sqrt{(\omega/c_2)^2-k_x^2}$$
$$\beta=\sqrt{l\Omega^2-\zeta^2}=\frac{2h}{\pi}\sqrt{(\omega/c_1)^2-k_x^2}$$
$$l=\frac{1-2\nu}{2(1-\nu)} \tag{5}$$

Here ω is the frequency; k_x is the wave number in the x direction; $C_1, \ldots, C_4$ are arbitrary constants; ζ is the nondimensional wave number in the x direction; and α and β are nondimensional wave numbers in the y direction for SV and P waves, respectively; ν is Poisson's ratio.

The expressions for displacements and the two needed components of stress are

$$u=\frac{i\pi}{2h}\left(\zeta C_1\,e^{i\pi\beta y_N/2h}+\zeta C_2\,e^{-i\pi\beta y_N/2h}\right.$$
$$\left.+\alpha C_3\,e^{i\pi\alpha y_N/2h}-\alpha C_4\,e^{-i\pi\alpha y_N/2h}\right)\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right]$$
$$v=\frac{i\pi}{2h}\left(\beta C_1\,e^{i\pi\beta y_N/2h}+\beta C_2\,e^{-i\pi\beta y_N/2h}\right.$$
$$\left.-\zeta C_3\,e^{i\pi\alpha y_N/2h}-\zeta C_4\,e^{-i\pi\alpha y_N/2h}\right)\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right]$$
$$\tau_{xy}=\frac{-\pi^2\mu}{4h^2}\left(2C_1\beta\zeta\,e^{i\pi\beta y_N/2h}-2C_2\beta\zeta e^{-i\pi\beta y_N/2h}\right.$$
$$\left.+C_3\varphi_1\,e^{i\pi\alpha y_N/2h}+C_4\varphi_1\,e^{-i\pi\alpha y_N/2h}\right)\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right]$$
$$\tau_{yy}=\frac{-\pi^2\mu}{4h^2}\left(C_1\varphi_1\,e^{i\pi\beta y_N/2h}+C_2\varphi_1\,e^{-i\pi\beta y_N/2h}\right.$$
$$\left.-2\alpha\zeta C_3\,e^{i\pi\alpha y_N/2h}+2\alpha\zeta C_4\,e^{-i\pi\alpha y_N/2h}\right)\exp\left[\frac{i\pi}{2h}(\zeta x-c_2\Omega t)\right] \tag{6}$$

where $\varphi_1=\alpha^2-\zeta^2$.

A similar procedure for the fast layer produces the corresponding solution for that layer. The expressions for the displacements and tractions are similar to those derived in the foregoing with the corresponding primed parameters replacing the unprimed ones. The continuity of the displacements and tractions across the interface of the two layers within a cell and across the interface of the two adjacent cells supply eight equations. Periodicity of the coefficients of the wave equations enable us to make use of the Floquet theorem and relate the solution in two adjacent cells. In this way the eight equations can be written only in terms of parameters of the Nth unit cell. These are eight equations in eight unknowns, which, in this case, are the arbitrary constants appearing in the general solution to the wave equations in each layer. They may be written in matrix notation as

$$\begin{bmatrix}
\zeta B_- & \zeta B_+ & \alpha A_- & -\alpha A_+ & -\zeta B'_+ & -\zeta B'_- & -\alpha' A'_+ & \alpha' A'_- \\
\beta B_- & -\beta B_+ & -\zeta A_- & -\zeta A_+ & -\beta' B'_+ & \beta' B'_- & \zeta A'_+ & \zeta A'_- \\
2\gamma\beta\zeta B_- & -2\gamma\beta\zeta B_+ & \gamma\varphi A_- & \gamma\varphi A_+ & -2\beta'\zeta B'_+ & 2\beta'\zeta B'_- & -\varphi' A'_+ & -\varphi' A'_- \\
\gamma\varphi B_- & \gamma\varphi B_+ & -2\gamma\alpha\zeta A_- & 2\gamma\alpha\zeta A_+ & -\varphi' B'_+ & -\varphi' B'_- & 2\alpha'\zeta A'_+ & -2\alpha'\zeta A'_- \\
\zeta B_+ & \zeta B_- & \alpha A_+ & -\alpha A_- & -\zeta B'_- \tau & -\zeta B'_+ \tau & -\alpha' A'_- \tau & \alpha' A'_+ \tau \\
\beta B_+ & -\beta B_- & -\zeta A_+ & -\zeta A_- & -\beta' B'_- \tau & \beta' B'_+ \tau & \zeta A'_- \tau & \zeta A'_+ \tau \\
2\gamma\beta\zeta B_+ & -2\gamma\beta\zeta B_- & \gamma\varphi A_+ & \gamma\varphi A_- & -2\beta'\zeta B'_- \tau & 2\beta'\zeta B'_+ \tau & -\varphi' A'_- \tau & -\varphi' A'_+ \tau \\
\gamma\varphi B_+ & \gamma\varphi B_- & -2\gamma\alpha\zeta A_+ & 2\gamma\alpha\zeta A_- & -\varphi' B'_- \tau & -\varphi' B'_+ \tau & 2\alpha'\zeta A'_- \tau & -2\alpha'\zeta A'_+ \tau
\end{bmatrix}
\begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C'_1 \\ C'_2 \\ C'_3 \\ C'_4 \end{bmatrix} = 0 \tag{7}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$A_{\pm} = \exp[\pm i\pi\alpha/2]$$
$$B_{\pm} = \exp[\pm i\beta/2]$$
$$\varphi = \alpha^2 - \zeta^2$$
$$\tau = \exp[ip\eta]$$
$$A_{\pm} = \exp[\pm i\pi\alpha/2]$$
$$B_{\pm} = \exp[\pm i\pi\beta/2]$$
$$\varphi' = \alpha'^2 - \zeta^2$$
$$\epsilon = \frac{h'}{h}$$
$$\gamma = \frac{\mu}{\mu'}$$
$$\sigma^2 = \frac{\mu\rho'}{\mu'\rho} = \frac{c_2^2}{c_2'^2}$$
$$l' = \frac{1-2\nu'}{2(1-\nu')}$$
$$\eta = \frac{2h}{\pi} k_y$$
$$p = \pi(1+\epsilon)$$
$$\alpha' = \sqrt{\sigma^2\Omega^2 - \zeta^2} = \frac{2h}{\pi} \sqrt{(\omega/c_2')^2 - k_x^2}$$
$$\beta' = \sqrt{\sigma^2\Omega^2 l' - \zeta^2} = \frac{2h}{\pi} \sqrt{(\omega/c_1')^2 - k_x^2} \tag{8}$$

where k_y is the Floquet wave number.

For a nontrivial solution to exist, the determinant of the system must be zero, thus leading to the dispersion equation. This determinant has been expanded and is given in closed form in reference [7] as

$$\begin{aligned}
&2\tau^2[-L_1(c_\alpha c_\beta + c_{\alpha'} c_{\beta'}) + L_2 s_\alpha s_\beta (1 - c_{\alpha'} c_{\beta'}) \\
&+ L_3 s_{\alpha'} s_{\beta'} (1 - c_\alpha c_\beta) - L_5 c_\alpha s_\beta c_{\alpha'} s_{\beta'} - L_6 s_\alpha c_\beta s_{\alpha'} c_{\beta'} \\
&- L_8 c_\alpha s_\beta s_{\alpha'} c_{\beta'} - L_9 s_\alpha c_\beta c_{\alpha'} s_{\beta'} + (L_{11} + L_{12}) c_\alpha c_\beta c_{\alpha'} c_{\beta'} \\
&+ \frac{1}{2} L_{10} s_\alpha s_\beta s_{\alpha'} s_{\beta'}] + \tau(\tau^2+1) \times [-L_4(c_\alpha c_{\alpha'} + c_\beta c_{\beta'}) \\
&+ L_5 s_\beta s_{\beta'} + L_6 s_\alpha s_{\alpha'} + L_7(c_\alpha c_{\beta'} + c_\beta c_{\alpha'}) + L_8 s_\beta s_{\alpha'} \\
&+ L_9 s_\alpha s_{\beta'}] + 2\tau^2 L_{11} + (\tau^4+1) L_{12} = 0
\end{aligned} \tag{9}$$

where

$$\begin{aligned}
c_\alpha &= \cos \pi\alpha \\
c_\beta &= \cos \pi\beta \\
c_{\alpha'} &= \cos \pi\alpha'\epsilon \\
c_{\beta'} &= \cos \pi\beta'\epsilon \\
s_\alpha &= \sin \pi\alpha \\
s_\beta &= \sin \pi\beta \\
s_{\alpha'} &= \sin \pi\alpha'\epsilon \\
s_{\beta'} &= \sin \pi\beta'\epsilon
\end{aligned} \tag{10}$$

$$\begin{aligned}
L_1 &= 2\alpha\beta\alpha'\beta' P_1 P_2 P_4 P_5 \\
L_2 &= \alpha'\beta'[(\alpha\beta P_1 P_2)^2 + (P_4 P_5)^2] \\
L_3 &= \alpha\beta[(P_1 P_4)^2 + (\alpha'\beta' P_2 P_5)^2] \\
L_4 &= 2\alpha\beta\alpha'\beta' P_1 P_5 P_3 P_6 \\
L_5 &= \alpha\alpha' P_3 P_6[(P_1\beta)^2 + (P_5\beta')^2] \\
L_6 &= \beta\beta' P_3 P_6[(P_1\alpha)^2 + (P_5\alpha')^2] \\
L_7 &= 2\alpha\beta\alpha'\beta' P_3 P_6 P_2 P_4 \\
L_8 &= \alpha\beta' P_3 P_6[(P_2\beta\alpha')^2 + P_4^2] \\
L_9 &= \beta\alpha' P_3 P_6[(P_2\alpha\beta')^2 + P_4^2] \\
L_{10} &= (\alpha\beta)^2 P_1^4 + (\alpha\beta\alpha'\beta')^2 P_2^4 + P_4^4 + (\alpha'\beta')^2 P_5^4 \\
&\quad + (P_3 P_6)^2 [(\alpha\beta')^2 + (\alpha'\beta)^2] \\
L_{11} &= \alpha\beta\alpha'\beta'[(P_1 P_5)^2 + (P_2 P_4)^2] \\
L_{12} &= \alpha\beta\alpha'\beta'(P_3 P_6)^2
\end{aligned} \tag{11}$$

and

$$\begin{aligned}
P_1 &= \sigma^2\Omega^2 + 2\zeta^2(\gamma-1) \\
P_2 &= -2\zeta(\gamma-1) \\
P_3 &= \gamma\Omega^2 \\
P_4 &= \zeta[2\zeta^2(\gamma-1) + \Omega^2(\sigma^2-\gamma)] \\
P_5 &= \gamma\Omega^2 - 2\zeta^2(\gamma-1) \\
P_6 &= \sigma^2\Omega^2 \\
\tau^2 &= \cos 2p\eta + i \sin 2p\eta \\
\tau^3 + \tau &= 2\cos p\eta(\cos 2p\eta + i \sin 2p\eta) \\
\tau^4 + 1 &= 2\cos 2p\eta(\cos 2p\eta + i \sin 2p\eta) \\
\tau^2 + 1 &= 2\cos p\eta(\cos p\eta + i \sin p\eta) \\
(\tau^2+1)^2 &= 4\cos^2 p\eta(\cos 2p\eta + i \sin 2p\eta)
\end{aligned} \tag{12}$$

Discussion of the Dispersion Equation

Equation (9) is the dispersion equation for the propagation of harmonic waves in a periodically layered infinite medium, which relates the frequency Ω to the two wave numbers ζ and η. It is a fourth-order equation in $\tau = e^{ip\eta}$ and therefore has four roots in τ. As can be seen from this relation the roots for η are periodic, such that if η is a root, $\eta + 2n/(1+\epsilon)$ ($n = 1, 2, 3, \ldots$) is also a root. The basic period is, therefore, $2/(1+\epsilon)$. The four nondimensional Floquet wave numbers obtained for a pair of Ω, ζ can be categorized as

(*i*) $\eta_1, -\eta_1, \eta_2, -\eta_2$
η_1, η_2 real numbers
$0 < \eta_1 \le 1/(1+\epsilon), 0 < \eta_2 \le 1/(1+\epsilon)$

(*ii*) $\eta_1, -\eta_1, \eta_2, \bar{\eta}_2$
η_1 real number, η_2 complex number
$0 \le \eta_1 \le 1/(1+\epsilon)$
real part of $\eta_2 = 0$ or $1/(1+\epsilon)$

(*iii*) $\eta_1, \bar{\eta}_1, \eta_2, \bar{\eta}_2$
η_1, η_2 complex numbers
real part of $\eta_1 = 0$ or $1/(1+\epsilon)$
real part of $\eta_2 = 0$ or $1/(1+\epsilon)$

(*iv*) $\eta, -\eta, \bar{\eta}, -\bar{\eta}$
η complex
$0 <$ real part of $\eta \le 1/(1+\epsilon)$ (13)

A bar designates the complex conjugate quantity.

Consider a harmonic wave traveling in the layered medium in the state of plane strain with frequency Ω and let the x component of the wave number be ζ. Equation (9) then gives the component of the wave number in the direction normal to the layers (the Floquet wave number η). There are four Floquet wave numbers for a pair of Ω, ζ numbers. In the case of wave propagation in a homogeneous medium it is known

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

that for a pair of Ω, ζ numbers there are four wave number components in the y direction, i.e., α, $-\alpha$, β, $-\beta$. This shows the similarity between the two media with regard to harmonic wave propagation.

Reflection of Waves at a Free Boundary

To study the propagation of harmonic waves in a periodically layered half space we must derive the general expressions for τ_{xy}, τ_{yy} in infinite space and set them equal to zero at some plane parallel to the layering. Imagine that for a pair of numbers Ω, ζ, (equation (9)) is solved and the four wave numbers in the direction normal to the layers are obtained. The system (7) can be solved for each triplet $\Omega - \zeta - \eta$ and the expressions for τ_{xy}, τ_{yy} can derived.

The general expressions for τ_{xy} and τ_{yy} will be the superposition of the four solutions, each corresponding to one Floquet wave number. In the soft layer we obtain

$$\tau_{xy} = (Cf_1 + Df_3 + Ef_5 + Ff_7) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right]$$

$$\tau_{yy} = (Cf_2 + Df_4 + Ef_6 + Ff_8) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right] \quad (14)$$

where C, D, E, and F are arbitrary constants and $f_j = f_j(\Omega, \zeta, y_N)$, $j = 1, 8$. In the stiff layer f_j must be replaced by f_j'.

The expressions for tractions at a cell a distance nd ($n = 1, 2, 3, \ldots$) away from the cell in which the solution was obtained are, considering the Floquet theorem,

$$\tau_{xy} = \left(Cf_1\, e^{ip\eta_1 n} + Df_3\, e^{-ip\eta_1 n} + Ef_5\, e^{ip\eta_2 n} + Ff_7\, e^{-ip\eta_2 n}\right) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right]$$

$$\tau_{yy} = \left(Cf_2\, e^{ip\eta_1 n} + Df_4\, e^{-ip\eta_1 n} + Ef_6\, e^{ip\eta_2 n} + Ff_8\, e^{-ip\eta_2 n}\right) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right] \quad (15)$$

The expressions for τ_{xy}' and τ_{yy}' can be obtained for replacing f_j by f_j'. Thus having the solution in one cell is tantamount to having the solution in all cells.

Imagine now that at a certain location $y_N = y_N^*$ in the slow layer the tractions vanish

$$Cf_1^* + Df_3^* + Ef_5^* + Ff_7^* = 0$$
$$Cf_2^* + Df_4^* + Ef_6^* + Ff_8^* = 0 \quad (16)$$

where $f_j^* = f_j(\Omega, \zeta, y_N^*)$, $j = 1, 8$. By setting $E = 0$ the following wave system can be envisaged. An incident wave with the y component of the wave number η_1 and amplitude C is reflected as two waves having η_1, η_2 as wave number components in the y direction and D and F as their amplitudes, respectively.

$$Cf_1^* + Df_3^* + Ff_7^* = 0$$
$$Cf_2^* + Df_4^* + Ff_8^* = 0$$
$$D = -\frac{f_1^* f_8^* - f_7^* f_2^*}{f_3^* f_8^* - f_7^* f_4^*} C$$
$$F = -\frac{f_3^* f_2^* - f_1^* f_4^*}{f_3^* f_8^* - f_7^* f_4^*} C \quad (17)$$

By setting $C = 0$, another system of waves can be envisaged which can be similarly described. An incident wave with y component of the wave number η_2 and amplitude E is reflected as two waves having η_1, u_2 as y components of the wave numbers and D and F as amplitudes, respectively.

$$Df_3^* + Ff_7^* = -Ef_5^*$$
$$Df_4^* + Ff_8^* = -Ef_6^*$$
$$D = -\frac{f_5^* f_8^* - f_6^* f_7^*}{f_3^* f_8^* - f_7^* f_4^*} E$$
$$F = -\frac{f_3^* f_6^* - f_5^* f_4^*}{f_3^* f_8^* - f_7^* f_4^*} E \quad (18)$$

Consider the equation

$$f_3^* f_8^* - f_7^* f_4^* = 0 \quad (19)$$

The expression on the left-hand side of this equation is the determinant of the system of equation (17) or (18). For a solution to exist it must be nonzero. If it is zero, however, by setting C and E equal to zero in the systems (17) and (18), respectively, we obtain the following homogeneous system of equations

$$Df_3^* + Ff_7^* = 0$$
$$Df_4^* + Ff_8^* = 0 \quad (20)$$
$$F = -\frac{f_3^*}{f_7^*} D$$
$$\tau_{xy} = (Df_3 + Ff_7) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right] \quad (21)$$
$$\tau_{yy} = (Df_4 + Ff_8) \exp\left[\frac{i\pi}{2h}(\zeta x - c_2\Omega t)\right]$$

For the wave represented by the preceding system to decay away from the surface, both η_1 and η_2 must have imaginary parts. It has been found that in the domains of the $\Omega - \zeta$ plane where equation (19) is valid, the foregoing condition is satisfied, which means system (20) represents a surface wave traveling in the layered medium. Equation (19) is the dispersion relations for that wave.

Types of Waves in a Half Space

In the case of harmonic wave propagation in the homogeneous half space, it is noted that the $\Omega - \zeta$ plane can be divided into three regions, in each of which certain types of waves travel. The regions are separated by the P line ($\beta = 0$) and by the SV line ($\alpha = 0$). A similar division can be introduced in the $\Omega - \zeta$ plane for the layered half space. Here the lines of division are given by $\eta_1 = 0$, $\eta_1 = 1/(1 + \epsilon)$, $\eta_2 = 0$, $\eta_2 = 1/(1 + \epsilon)$. These curves represent wave motions in the infinite layered medium parallel to the layers. The four types of such curves correspond to the symmetry or antisymmetry with respect to the centerline of the displacements u and u' in the soft or fast layers, respectively. With respect to the roots of equation (7) region *I* corresponds to the category (*i*) of roots, region *II* corresponds to the category (*ii*), and region *III* corresponds to categories (*iii*) and (*iv*). Region *IIIa* corresponds to category (*iii*) of the roots of equation (7) and region *IIIb* corresponds to category (*iv*) of these roots. As can be seen from Fig. 2 between the boundaries of region *IIIb* and *II* there may be regions that correspond to category (*i*) of roots. On the boundaries of region *IIIb* the two Floquet wave numbers are equal.

The computation was carried out for the flowing values of the parameters

$$\gamma = 0.02$$
$$\sigma^2 = 0.06$$
$$\epsilon = 4$$
$$\nu = 0.3$$
$$\nu' = 0.35$$

In the layered half space for points in region *I* (Fig. 2), both Floquet wave numbers η_1 and η_2 are real which means

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

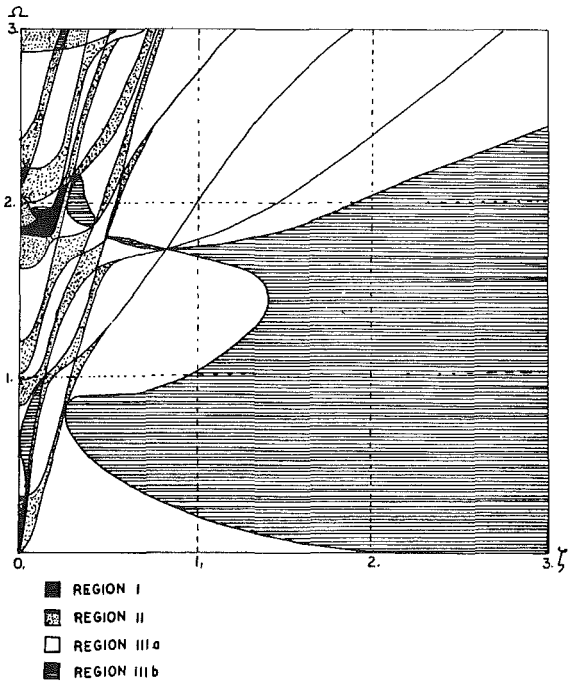


Fig. 2 Regions of $\Omega - \zeta$ plane

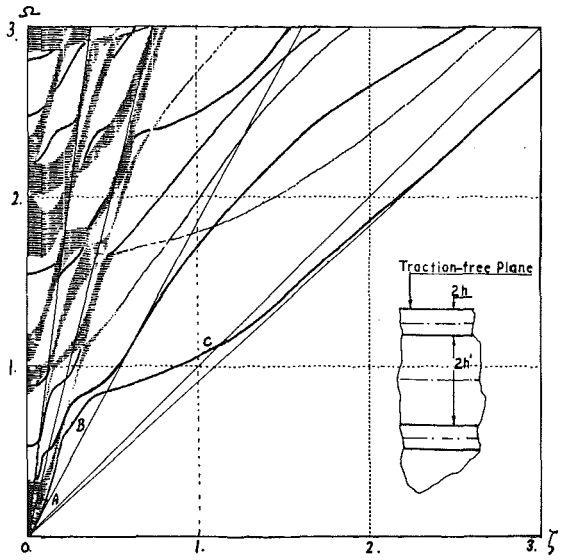


Fig. 3 Dispersion curve for surface waves. Slow layer at the surface.

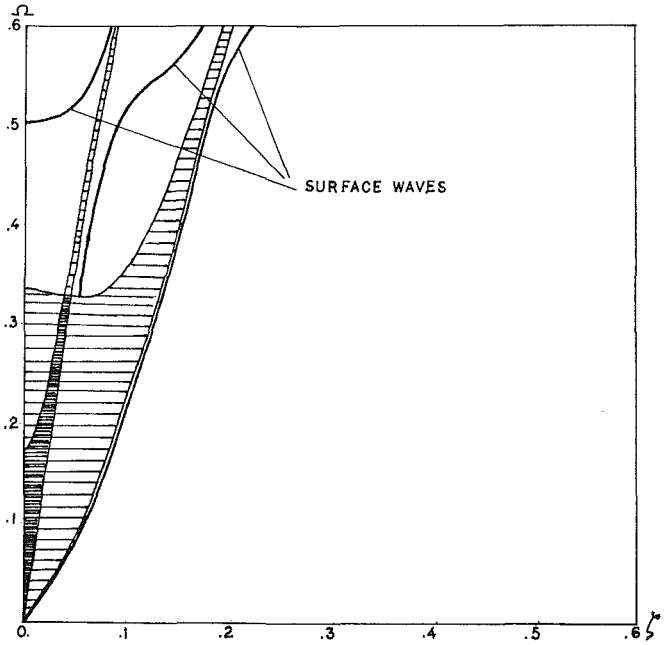


Fig. 4 Dispersion curve for surface waves. Magnified view of part of Fig. 3.

propagation of bulk waves in the half space. For points in region *II* either η_1 or η_2 has an imaginary part, which means an incident bulk wave reflects as a bulk wave and as a surface wave. For points in region *III* both η_1 and η_2 have imaginary parts. Therefore, an incident wave grows away from the surface while the two reflected waves decay. The reflection patterns described in the foregoing are governed by systems of equations (17) or (18).

It is only in region *III*, where the two reflected waves decay that we expect the determinant of the systems (17) or (18) to vanish along some curves so that a system of surface waves may result, similar to Rayleigh waves in a homogeneous medium. As seen from Fig. 2, the regions II and *III* have multiple boundaries and since surface waves can exist only in region *III* we expect the dispersion curve for such surface waves to start and end on the boundaries of the two regions.

Surface Waves in Layered Medium

The system of equations (17) and (18) is solvable if the determinant of the coefficients does not vanish. However, if it vanishes the situation is similar to the vanishing of the determinant of systems in a homogeneous medium which resulted in Rayleigh waves. The Rayleigh line represents points along which the determinant vanishes.

In the present investigation curves were sought along which a similar phenomenon occurs in the layered half space. Equation (19) was solved numerically and the dispersion curve thus obtained is shown in Fig. 3, and on a magnified scale in Fig. 4. The dispersion curve shown in this figure is for a traction-free plane at $y_N = -h$.

As can be seen from Fig. 3, there is an infinite number of branches for the dispersion curve. Whenever the dispersion branches reach region *II* they come to an end. The determinant is then no longer zero and the system of waves governed by equations (17) or (18) prevails. As one moves along a dispersion branch, the closer one gets to the end points the weaker is the decay rate and the deeper does the wave penetrate into the half space. At the end point itself a transition from surface waves to bulk waves occurs as region *II* is reached.

For $\zeta = 0$ the system of plane strain waves traveling in the layered half space and described by equation (7) decouples into its constituent parts, i.e., a *P* and as *SV* wave traveling normal to the layers. At such frequencies the traction-free surface repeats itself periodically. Therefore there is an infinite number of traction-free surfaces in contrast to the general case for nonzero ζ where there is only one traction-free surface.

As the frequency Ω and the wave number ζ tend to zero the system of equations (7) must be solved in the vicinity of the origin. The phase velocity (which is the slope of the tangent to the dispersion curve at the origin) computed in this manner agrees with the one predicted by the effective modulus theory for a homogeneous, transversely isotropic medium [8-9]. For the parameters chosen in this paper, the slope at the origin is $\Omega/\zeta = 1.32$. As ζ becomes large, depending on the location of the free surface, the lowest branch will tend to the Rayleigh line for the layer in which the free surface resides. As mentioned, Fig. 3 is for the case when the slow layer is the bounding top layer. Results for calculations when the fast layer is the bounding top layer are not given here, but may be found in reference [10].

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

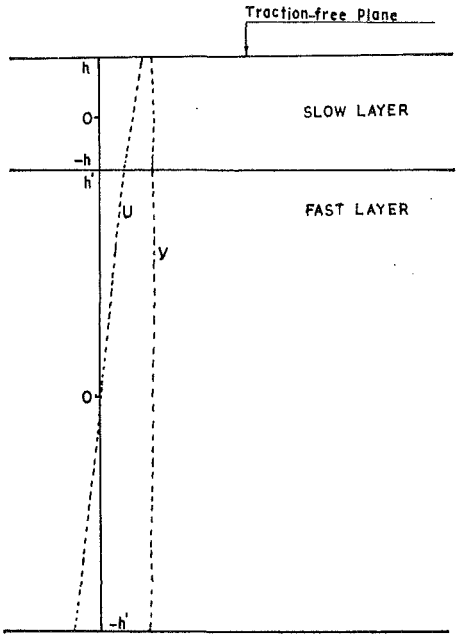


Fig. 5 Horizontal and vertical displacements of the first cell for Ω = 0.2 (point *A* in Fig. 3)

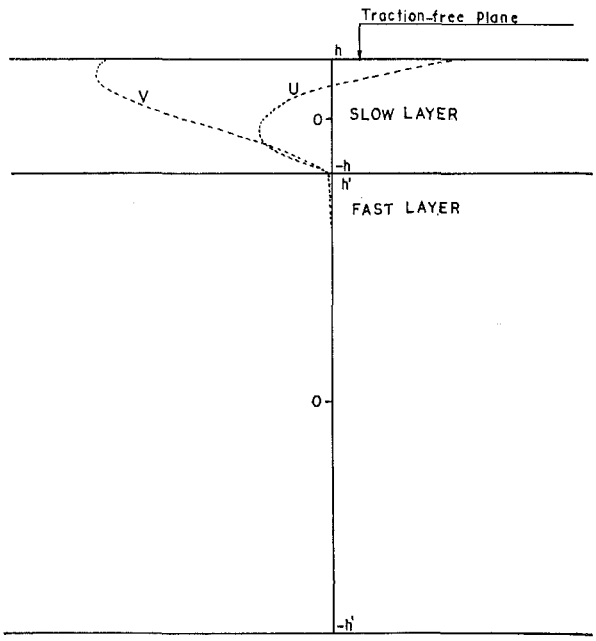


Fig. 7 Horizontal and vertical displacements of the first cell for Ω = 1.2 (point *C* in Fig. 3)

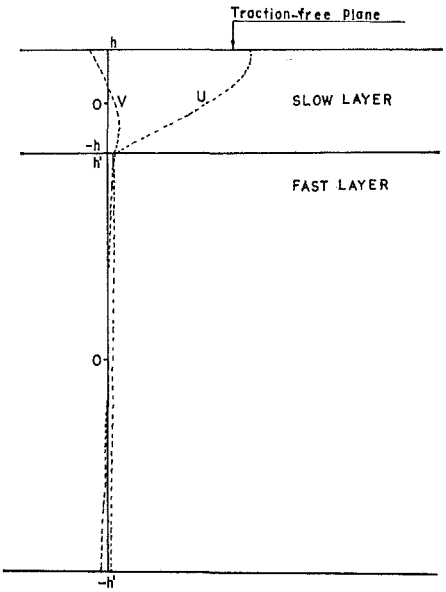


Fig. 6 Horizontal and vertical displacements of the first cell for Ω = 0.65 (point *B* in Fig. 3)

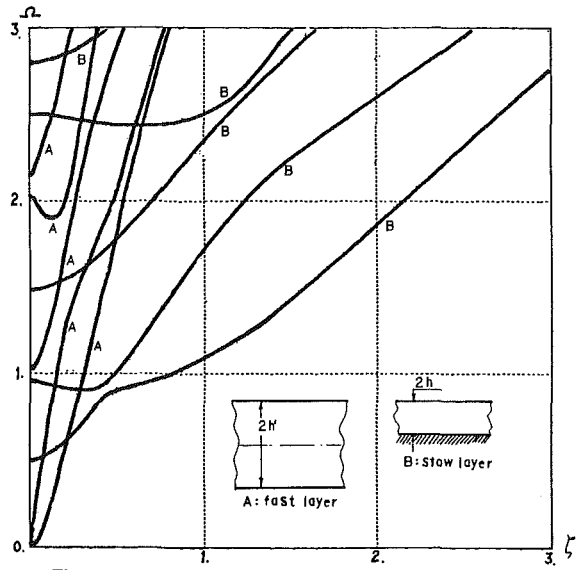


Fig. 8 The superposition of dispersion curve of clamped-free, slow layer and free-free, fast layer

Analysis of the Lowest Surface Wave Branch

To gain some insight into the features of the lowest branch of the surface wave, whose dispersion is given in Fig. 3, it is helpful to calculate the mode shapes and attempt to obtain the reasons for the dispersive behavior. To begin with, it is observed from Fig. 3 that this lowest branch is continuous in the complete range $0 \leq \xi \leq \infty$ and is composed essentially of three segments. Typical points on each of the three segments are denoted by *A*, *B*, and *C*, respectively. Mode shapes for the top two layers, i.e., the top unit cell, are given in Figs. 5–7, for points *A*, *B*, and *C*, respectively. It is noted that at point *A* the fast layer appears to undergo simple bending, while at points *B* and *C* all motion is essentially confined to the top slow layer and the fast layer remains essentially undisplaced.

These properties of modes of the lowest branch of the surface wave suggest an analysis of dispersive properties of two much simpler systems in the expectation that these would be similar to those of the layered half space under discussion in the present paper. The two simple systems consists of: (*A*) a plate with traction-free faces, thickness $2h'$, and properties of the fast layer; (*B*) a plate of thickness $2h$, one face traction-free, the other face clamped, and the properties of the slow layer. Dispersion curves for these two plates are sketched in Fig. 8. It is noted that the lowest continuous curve, consisting, as ξ increases, first of the lowest curve *A* and then of the lowest curve *B*, already greatly resembles the lowest dispersion curve of the layered half space, Fig. 3. The behavior near the origin is represented in magnified form in Fig. 9. Here the second mode of system *A*, namely, the lowest symmetric mode, as well as the second and third branches of the surface wave in the semi-infinite layered medium have also been included.

This figure illustrates not only the closeness of the lowest

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

surface wave branch and the lowest asymmetric (bending) mode of the free-free stiff plates for low frequencies up to the first cut-off frequency of a surface wave branch (occurring for the third-lowest branch), but also the feature that the second lowest and the third lowest branches of the surface wave are uncoupled by the lowest symmetric mode of the free-free fast plate and the lowest mode of the clamped-free slow plate.

Returning to Fig. 8 it is observed that the lowest curve B itself appears to consist of two parts separated by the approximate value of $\zeta = 0.5$. It is thus tempting to investigate more closely the properties of those two regimes. With this aim in mind it is noted that system B, i.e., clamped-free layer of thickness $2h$, inherently lacks symmetry, and thus its modes cannot be grouped into symmetric and asymmetric ones. To alleviate this undesirable feature we consider a substitute system consisting of a slow layer of thickness $4h$ clamped at both faces. This permits one to consider the lowest symmmetric and the lowest asymmetric modes of this plate, whose dispersion is plotted as curves 2 and 3 in Fig. 10. The lowest asymmetric mode of the free-free stiff plate is replotted in this figure, as well as the Rayleigh line for the slow material. It is observed that curves 2 and 3 represent, to a good accuracy, the uncoupled dispersion curves of the two lowest modes of system B of Fig. 8, as given by the two lowest curves in that figure.

The lowest dispersion curve for surface waves in the layered medium has also been reproduced in Fig. 10 and it is seen that the desired result has been achieved to a larger extent: the behavior of the fundamental surface wave can now be physically understood over the whole range of wave number. For very long waves and while the wave length is larger than approximately twice the cell thickness $d = 2h + 2h'$, the system behaves essentially as a free-free plate of thickness $2h'$. If the wave length is reduced further, to approximately the cell thickness d, the system is in the intermediate regime denoted by point B in Fig. 3 as discussed in the foregoing. For small wave lengths the system behaves first as one of the lowest modes of system B and then asymptotically approaches the Rayleigh line of the slow, top layer.

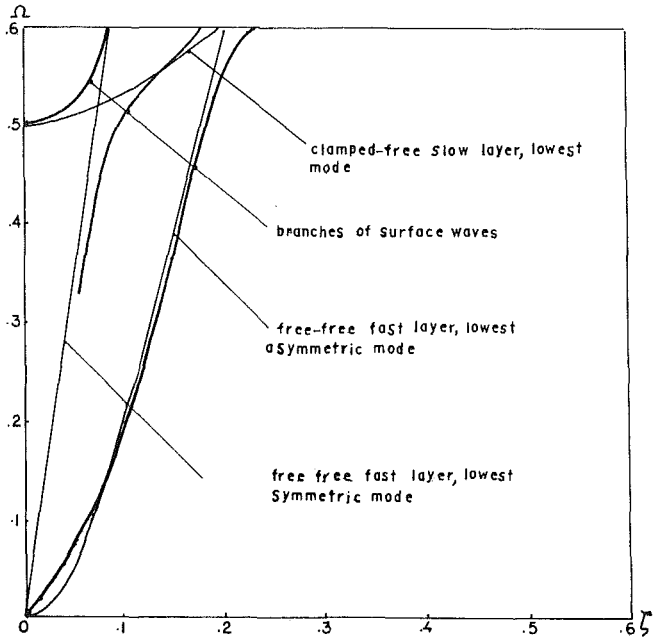


Fig. 9 Surface waves dispersion branches and uncoupling branches of a clamped-free, slow layer and free-free, fast layer

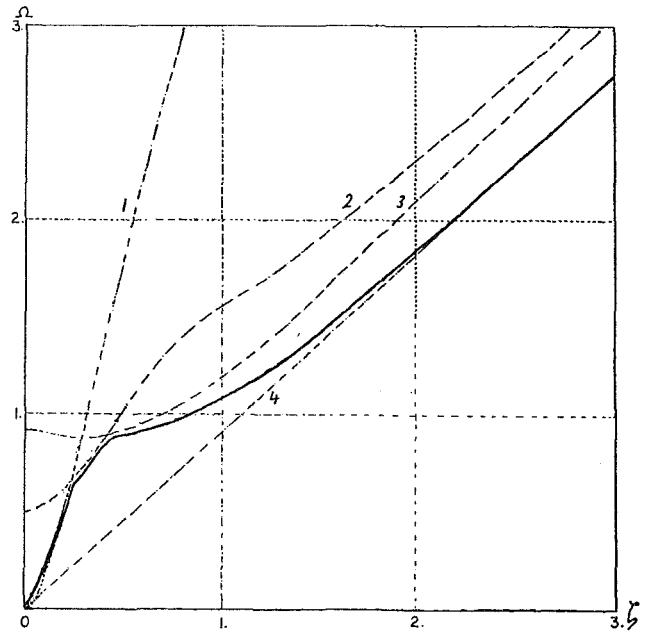


Fig. 10 The lowest branch of the dispersion curve with the fast layer at the surface and the uncoupling curves. 1—Lowest asymmetric branch of the free-free stiff plate. 2,3—Lowest branches of a clamped-clamped soft plate two times the thickness of the slow layer. 4—Rayleigh line for the slow medium.

Acknowledgment

This research was supported in part by ONR Contract N00014-76-C-0054 to Stanford University.

References

1 Auld, B. A., Beaupre, G. S., and Herrmann, G., "Horizontal Shear Surface Waves on a Laminated Composite," *Electronic Letters*, Vol. 13, 1977, pp. 525-527.

2 Auld, B. A., Beaupre, G. S., and Herrmann, G., "Horizontal Shear Surface Waves on a Laminated Composite," in *Modern Problems in Elastic Wave Propagation*, Wiley, New York, 1978.

3 Herrmann, G., Beaupre, G. S., and Auld, B. A., "Applicability of Floquet-Type Solutions to Bounded Layered Composites," in *Mechanics Today*, Vol. 5, Nemat-Nasser, S., ed., Pergamon Press, 1980, pp. 83-93.

4 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body, Antiplane Strain," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 343-349.

5 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body, Plane Strain," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 113-119.

6 Delph, T. J., Herrmann, G., and Kaul, R. K., "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body, Plane Strain, Numerical Results," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 531-537.

7 Golebiewska, A. A., "On Dispersion of Periodically Layered Composites in Plane Strain," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 206-207.

8 Postma, G. W., "Wave Propagation in a Stratified Medium," *Geophysics*, Vol. 20, 1955, pp. 780-806.

9 Deresiewicz, H., and Mindlin, R. D., "Waves on the Surface of a Crystal," *Journal of Applied Physics*, Vol. 28, 1957, pp. 669-671.

10 Hemami, M., "Wave Propagation in a Layered Half-Space," Ph.D. Dissertation, Stanford University, Stanford, Calif., Nov. 1980.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Plastic Stress Intensity Factors for Small Scale Yielding in Antiplane Shear by Caustics

P. S. Theocaris
Mem. ASME

C. I. Razem

Department of Theoretical
and Applied Mechanics,
The National Technical University of Athens,
5, Heroes of Polytechnion Avenue,
Zographou, Athens 624, Greece

The K_{III}-stress intensity factor in an edge-cracked plate submitted to antiplane shear may be evaluated by the reflected caustic created around the crack tip, provided that a purely elastic behavior exists at the crack tip [1]. For a work-hardening, elastic-plastic material, when stresses at the vicinity of the crack tip exceed the yield limit of the material, the new shape of caustic differs substantially from the corresponding shape of the elastic solution. In this paper the shape and size of the caustics created at the tip of the crack, when small-scale yielding is established in the vicinity of the crack tip, were studied, based on a closed-form solution introduced by Rice [2]. The plastic stress intensity factor may be evaluated from the dimensions of the plastic caustic. Experimental evidence with cracked plates made of opaque materials, like steel, corroborated the results of the theory.

Introduction

The determination of the stress intensity factor at the crack tip of a thin sheet submitted to any mode of deformation, when plasticity or nonlinear elasticity is preceding fracture, is of fundamental importance. Such a solution is a difficult and rather impossible task.

Hult and McClintock [3] considered the problem of an edge crack in a semi-infinite plane for load levels inducing plastic zones, comparable to the dimensions of the crack. A solution was also presented by Rice [4] for the problem of a sharp notch or a crack in a plane of finite width subjected to antiplane shear for an elastic perfectly plastic material. Rice [2], also considered the problem of a plate with a crack in antiplane shear for a work-hardening, elastic-plastic material.

Up to now the optical method of reflected caustics was succeeded to evaluate the stress intensity factor in antiplane shear when elasticity was dominating the stress field near the crack tip [1].

In this paper the method of reflected caustics was extended to evaluate stress intensity factors in K_{III}-mode, when small-scale yielding is developed around the crack tip. Using the equations developed by Rice [2], a practical procedure was developed, which combines details of the dominant singularity at the plastic region, like the plastic stress intensity factor, with the geometrical elements of the corresponding caustic surrounding the crack tip.

The Optical Method of Caustics

The method of reflected caustics, which was successfully applied up to now for the study of the singular behavior of elastic stress fields, may be extended to study singularities in the presence of plasticity at the neighborhood of the singular field. The method is based on a principle of transformation of the stress singularity to a respective optical singularity.

It has been shown [5] that the generatrix curve (*initial curve*), which is the locus on the surface of the specimen of reflections of light rays creating the caustic, is given by:

$$1+\left(\frac{2z_0}{\lambda}\right)^2(P+Q-R^2)+\frac{2z_0}{\lambda}(P+Q)=0 \tag{1}$$

where z_0 and λ denote the distance of the reference screen from the specimen and the magnification factor of the optical system. The coefficients P, Q, and R are given by:

$$P=\frac{\partial^2 w}{\partial r^2},\quad Q=\frac{1}{r^2}\frac{\partial^2 w}{\partial \vartheta^2}+\frac{1}{r}\frac{\partial w}{\partial r},\quad R=\frac{1}{r}\frac{\partial^2 w}{\partial r\partial \vartheta}-\frac{1}{r^2}\frac{\partial w}{\partial \vartheta} \tag{2}$$

where (r, ϑ) are the polar coordinates of a system with the origin at the crack tip and the Ox-axis coinciding with the crack axis (Fig. 1).

Moreover, the parametric equations of the caustic are given by:

$$X=\lambda r\cos\vartheta+2z_0\left(\frac{\partial w}{\partial r}\cos\vartheta-\frac{1}{r}\frac{\partial w}{\partial \vartheta}\sin\vartheta\right) \tag{3}$$

$$Y=\lambda r\sin\vartheta+2z_0\left(\frac{\partial w}{\partial r}\sin\vartheta+\frac{1}{r}\frac{\partial w}{\partial \vartheta}\cos\vartheta\right) \tag{4}$$

The shape and the dimensions of the curve determined by equations (1) - (4) depend on the stress field near the crack tip and they will be used for the experimental determination of the plastic stress intensity factor.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on the paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November, 1981; final revision, May, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

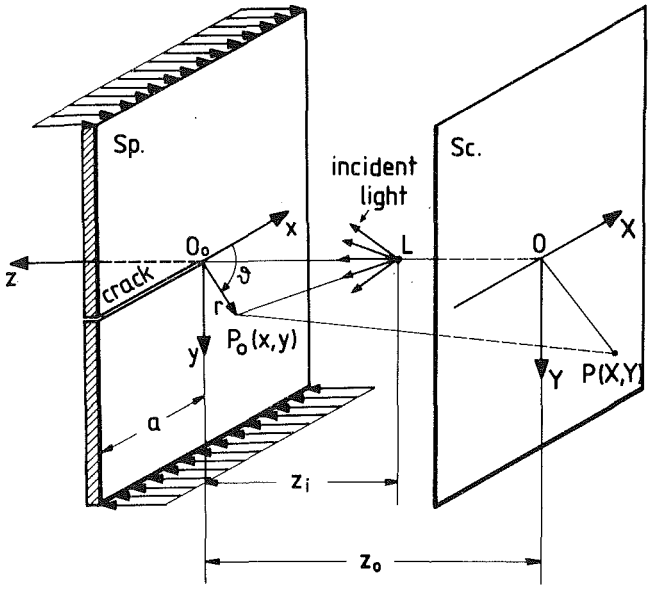


Fig. 1 Schematic diagram of the experimental set-up

Caustics at Crack Tips in Antiplane Elastic-Plastic Shear

Consider a semi-infinite orthogonal plate containing an edge crack of length a, submitted to an antiplane shear loading, as in Fig. 1. The dominant singular solution for a power-hardening material is given in the form [2,6]:

$$\bar{\tau}(\gamma) = \bar{\tau}_0 \left(\frac{\bar{\gamma}}{\bar{\gamma}_0} \right)^{1/n} \tag{5}$$

In this relation $\bar{\tau}_0$ and $\bar{\gamma}_0$ are the yield stress and strain in pure shear and $\bar{\tau}(\gamma)$ and $\bar{\gamma}$ are stresses and strains at a generic point in the vicinity of the crack tip, whereas n expresses the *power hardening coefficient*. In this relation, as well as throughout the paper, barred quantities represent physical quantities, whereas unbarred ones represent normalized quantities to the respective yield stress $\bar{\sigma}_0$ and strain $\bar{\gamma}_0$.

Introducing polar coordinates (r,ϑ) referred to the crack tip, as in Fig. 1, the solution for the deflections of the plate in the plastic zone may be derived in the form [7]:

$$w = K_\epsilon \left[\frac{h(\vartheta)}{r} \right]^{-1/(n+1)} \sin\rho \tag{6}$$

where:

$$2\rho = \vartheta - \arcsin\left[\left(\frac{n-1}{n+1} \right) \sin\vartheta \right] \tag{7}$$

$$h(\vartheta) = \frac{\sin 2\rho}{2\sin\vartheta} \tag{8}$$

$$K_\epsilon = (K_\sigma)^n \tag{9}$$

$$K_\sigma = (\tau_\infty)^{2/(n+1)} a^{1/(n+1)} \tag{10}$$

and the self-evident notations $\bar{K}_\sigma = \bar{K}_\sigma/(\bar{\tau}_0)$, $w = \bar{w}/\bar{\gamma}_0$, $\tau_\infty = \bar{\tau}_\infty/\bar{\tau}_0$, and $\bar{\gamma}_0 = \bar{\tau}_0/G$. In these relations, $\bar{w} = \bar{w}(r,\vartheta)$ is the z-displacement field of the specimen (deflections), $\bar{K}_\sigma$ is the plastic stress intensity factor, $\bar{\tau}_\infty$ and $\bar{\gamma}_\infty$ are the out-of-plane shearing stresses and strains at infinity, which are related by $\bar{\tau}_\infty = G\bar{\gamma}_\infty$ for an applied stress $\bar{\tau}_\infty$ below yielding.

Equation (6) expresses the full solution everywhere in the plastic zone [7]. The plastic zone is bounded by a curve, which, for small-scale yielding, becomes a circle of radius:

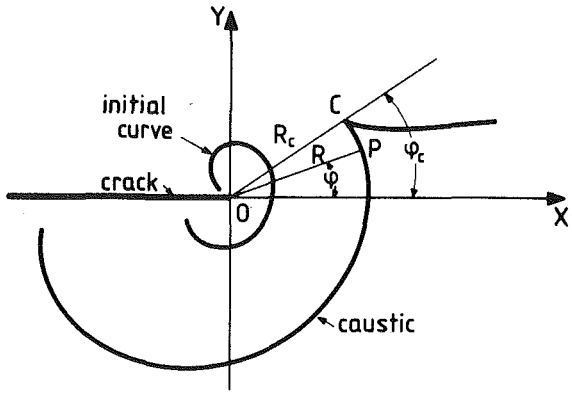


Fig. 2 Theoretically obtained caustic and its initial curve in plastic antiplane longitudinal shear

$$R_0 = \frac{a}{2} (\tau_\infty)^2 \tag{11}$$

with its center shifted along the crack axis at a distance

$$\beta_0 = (\tau_\infty)^2 \frac{a}{2} \frac{(n-1)}{(n+1)} \tag{12}$$

ahead of the crack tip [2].

Differentiating equation (6) with respect to r and ϑ successively and introducing these derivatives into relations (2) we can find from equation (1) the equation of the initial curve, expressed by:

$$1 + \left(\frac{2z_0}{\lambda} K_\epsilon r^{N-2} \right)^2 F_a(\vartheta,n) + \left(\frac{2z_0}{\lambda} K_\epsilon r^{N-2} \right) F_b(\vartheta,n) = 0 \tag{13}$$

In this relation $F_a(\vartheta, n)$ and $F_b(\vartheta, n)$ are expressed by:

$$F_a(\vartheta,n) = F_1(\vartheta,n)F_2(\vartheta,n) - [F_3(\vartheta,n)]^2 \tag{14}$$

$$F_b(\vartheta,n) = F_1(\vartheta,n) + F_2(\vartheta,n) \tag{15}$$

where the quantities F_1, F_2, and F_3 take the forms:

$$F_1(\vartheta,n) = N(N-1)h\sin\rho \tag{16}$$

$$\begin{aligned} F_2(\vartheta,n) = {} & N(N+1)h^{-N-2}\left(\frac{dh}{d\vartheta} \right)^2 \sin\rho \\ & - Nh^{-N-1} \frac{d^2h}{d\vartheta^2} \sin\rho - 2Nh^{-N-1} \frac{dh}{d\vartheta} \frac{d\rho}{d\vartheta} \cos\rho \\ & - h^{-N} \sin\rho \left(\frac{d\rho}{d\vartheta} \right)^2 + h^{-N} \frac{d^2\rho}{d\vartheta^2} \cos\rho + \\ & + Nh^{-N} \sin\rho \end{aligned} \tag{17}$$

$$F_3(\vartheta,n) = (1-N)\left[Nh^{-N-1} \frac{dh}{d\vartheta} \sin\rho - h^{-N} \frac{d\rho}{d\vartheta} \cos\rho \right] \tag{18}$$

Quantity N depends on the power-hardening coefficient of the material, n, and it is defined by:

$$N = \frac{1}{n+1} \tag{19}$$

When the functions $\rho \equiv \rho(\vartheta)$ and $h \equiv h(\vartheta)$, for angle ϑ varying between $-\pi$ and π, and their first and second derivatives are introduced into relation (13) they yield:

$$r_{1,2} = \left(\frac{2z_0}{\lambda} K_\epsilon \right)^{\frac{n+1}{2n+1}} [u_{1,2}(\vartheta,n)]^{-\frac{n+1}{2n+1}} \tag{20}$$

where the functions $u_{1,2}(\vartheta,n)$ are given by:

$$u_{1,2}(\vartheta,n) = \frac{-F_b(\vartheta,n) \pm ([F_b(\vartheta,n)]^2 - 4F_a(\vartheta,n))^{1/2}}{2F_a(\vartheta,n)} \tag{21}$$

The sign in front of the square root in relations (21) should

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

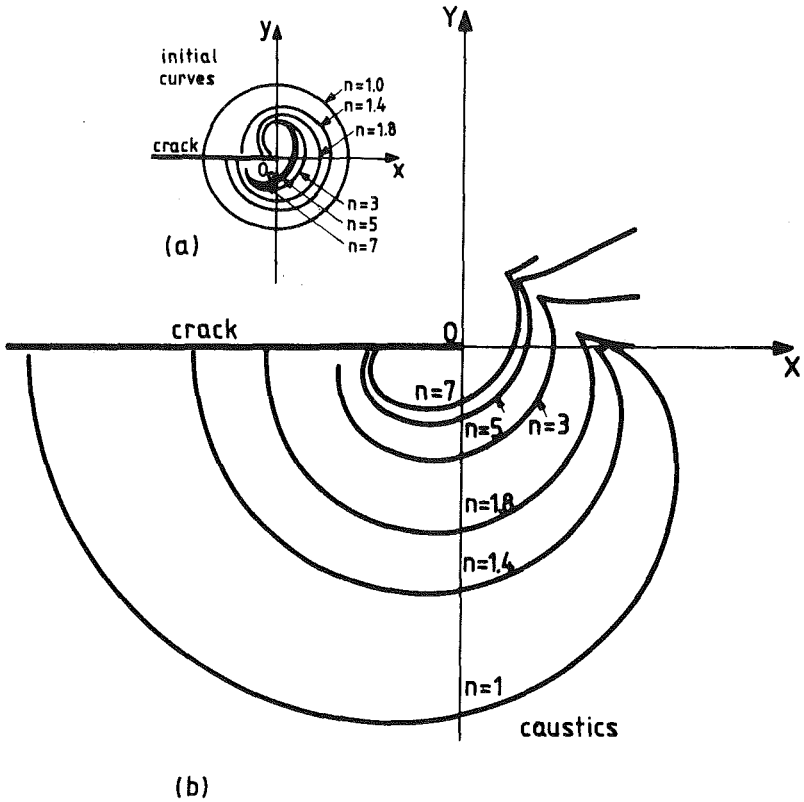


Fig. 3 Evolution of caustic and its initial curves from elastic ($n = 1$) to elastic-plastic ($n > 1$) state of stress near the crack tip

be selected in combination with the signs of K_ϵ and λ of relation (20) to yield a positive value for r. The factor λ is positive or negative, if the specimen is illuminated by a divergent or convergent light beam [5]. In our tests λ was always positive. The sign of the coefficient K_ϵ depends on the sign of the plastic stress intensity factor, K_σ. Moreover, the sign of the plastic stress intensity factor, K_σ, depends on the sign of the shear stress $\tilde{\tau}_\infty$ at infinity, which is always taken as positive. Finally, the sign of the expression (20) depends only on the signs of $u(\vartheta, n)$ and λ, whose product should always have a positive sign for the polar radius r of the initial curve.

Relation (20) indicates that the shape of the initial curve depends only on the value of the power-hardening coefficient n. It does not depend on the distance z_0, the magnification factor λ, and the plastic stress-intensity factor K_σ. On the other hand, the size of the initial curve depends on the value of the plastic stress intensity factor K_σ, the distance z_0, and the magnification factor λ of the optical set-up.

We now derive the parametric equations of the caustic. Equations (3) and (4) together with relations (6)–(10) yield:

$$X = \lambda r\left\{\cos\vartheta + \frac{2z_0}{\lambda}K_\epsilon r^{N-2}\left[Nh^{-N}\sin\rho\cos\vartheta + Nh^{-N-1}\frac{dh}{d\vartheta}\sin\rho\sin\vartheta - h^{-N}\frac{d\rho}{d\vartheta}\cos\rho\sin\vartheta\right]\right\} \tag{22}$$

$$Y = \lambda r\left\{\sin\vartheta + \frac{2z_0}{\lambda}K_\epsilon r^{N-2}\left[Nh^{-N}\sin\rho\sin\vartheta - Nh^{-N-1}\frac{dh}{d\vartheta}\sin\rho\cos\vartheta + h^{-N}\frac{d\rho}{d\vartheta}\cos\rho\cos\vartheta\right]\right\} \tag{23}$$

Taking into account equation (20) for the initial curve we can derive the parametric relations:

$$X = \lambda^{\frac{n}{2n+1}}(2z_0K_\epsilon)^{\frac{n+1}{2n+1}}F_x(\vartheta,n) \tag{24}$$

$$Y = \lambda^{\frac{n}{2n+1}}(2z_0K_\epsilon)^{\frac{n+1}{2n+1}}F_y(\vartheta,n) \tag{25}$$

In these relations the functions $F_x(\vartheta,n)$ and $F_y(\vartheta,n)$ are expressed by:

$$F_x(\vartheta,n) = (u(\vartheta,n))^{-\frac{n+1}{2n+1}}\left\{\cos\vartheta + u(\vartheta,n)\left[Nh^{-N}\sin\rho\cos\vartheta + Nh^{-N-1}\frac{dh}{d\vartheta}\sin\rho\sin\vartheta - h^{-N}\frac{d\rho}{d\vartheta}\cos\rho\cos\vartheta\right]\right\} \tag{26}$$

$$F_y(\vartheta,n) = (u(\vartheta,n))^{-\frac{n+1}{2n+1}}\left\{\sin\vartheta + u(\vartheta,n)\left[Nh^{-N}\sin\rho\sin\vartheta - Nh^{-N-1}\frac{dh}{d\vartheta}\sin\rho\cos\vartheta + h^{-N}\frac{d\rho}{d\vartheta}\cos\rho\cos\vartheta\right]\right\} \tag{27}$$

where the expressions $u(\vartheta,n)$ represent the solutions of equation (21) which satisfy relations (20) for a positive radius r of the initial curve.

It is clear from relations (24) and (25) that the shape of the caustic depends on the same quantities, as the shape of the initial curve. Therefore, the shape of the caustic depends only on the power-hardening coefficient n, whereas its size depends on the value of the plastic stress intensity factor K_σ, the magnification factor of the optical set-up λ, and on the distance z_0.

Figure 3(a,b) presents the shape of the theoretically obtained caustics and their initial curves for a strain-hardening, elastic-plastic stress field, corresponding to different values of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the power-hardening coefficient n. For the tracing of these caustics it was assumed that the reference screen is placed at a distance $z_0 = 1.50$ m from the plate and the magnification factor of the optical system is $\lambda = 4$. The elastic case corresponds to $n = 1$ [1].

When a plastic stress field appears around the crack tip (for $n > 1$), the caustic diminishes in size, changes its shape, and starts to present a cusp point. This cusp point is progressively displaced angularly about the crack tip. Moreover, a second branch of the caustic, on the other side of the cusp point, appears and it is continuously increasing, as n increases (Fig. 3(*b*)).

It has been shown [4] that the plastic zone, for mode-*III* deformation, is a circle, whose center lies in front of the crack tip and along the axis of the crack. Its periphery is either touching the crack tip for perfectly plastic materials, or cutting the lips of the crack. This geometry of the plastic zone implies that the main part of the initial curve, which lies in front of the crack, is always inside the plastic zone. However, for small-scale yielding the initial curve may lie only partially inside the plastic zone. Nevertheless, since the existence of a cusp point at the intersection of the two unequal branches of the caustic is characteristic only for a plastic zone around the crack tip, it is certain that the region surrounding this point always lies inside the plastic zone. During the plotting of the caustics the elastic-plastic boundary was not programmed to limit the tracing of the respective caustic only inside the plastic zone. Thus, the caustics traced in the figures of the paper, may have their tails lying outside the plastic zone. These extremities do not exist in reality and, indeed, these tails are missing in the experimentally determined caustics (compare Figs. 6(*b*) and 7 with Figs. 6(*a*) and 3). Although this limitation of the caustics on the respective elastic-plastic boundary could be programmed in the computer, it was avoided because of its unimportance and the complications introduced in the computer program.

The elastic lateral deformations along the thickness of the plate are much smaller than the plastic deformations creating the dimple in front of the crack [9]. Then, the variation of the

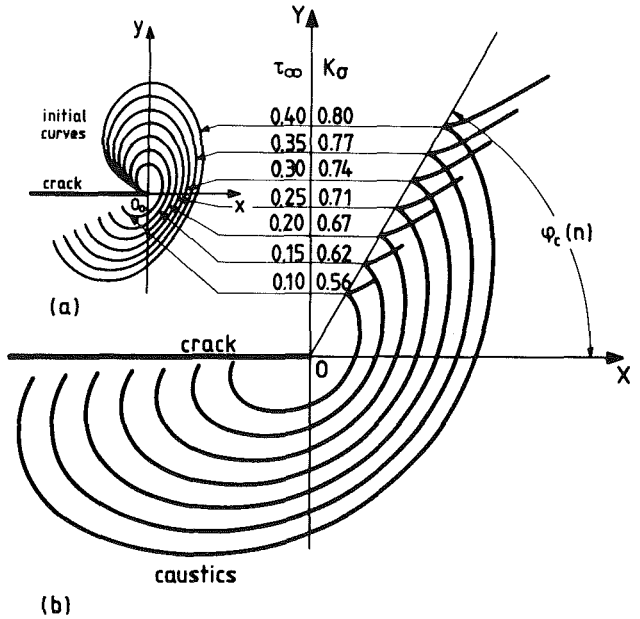


Fig. 4 Variation of the function $F(n)$ and polar angle of the cusp point $\varphi_c(n)$ versus the power-hardening coefficient n

Table 1 The values of the function $F(n)$ and the polar angle φ_c for the cusp point of the caustic for various values of the power-hardening coefficient

n	1	1.5	2	3	4	5	6	7
F	0.500	0.6417	0.6746	0.7379	0.8109	0.8731	0.9220	0.9602
φ_c	0.000	1.45°	13.09°	33.91°	45.80°	52.67°	57.13°	60.31°
n	8	9	10	11	12	13	14	15
F	0.9898	1.0130	1.0306	1.0454	1.0558	1.0645	1.0723	1.0800
φ_c	62.74°	64.64°	66.37°	67.72°	69.00°	70.11°	71.02°	71.83°

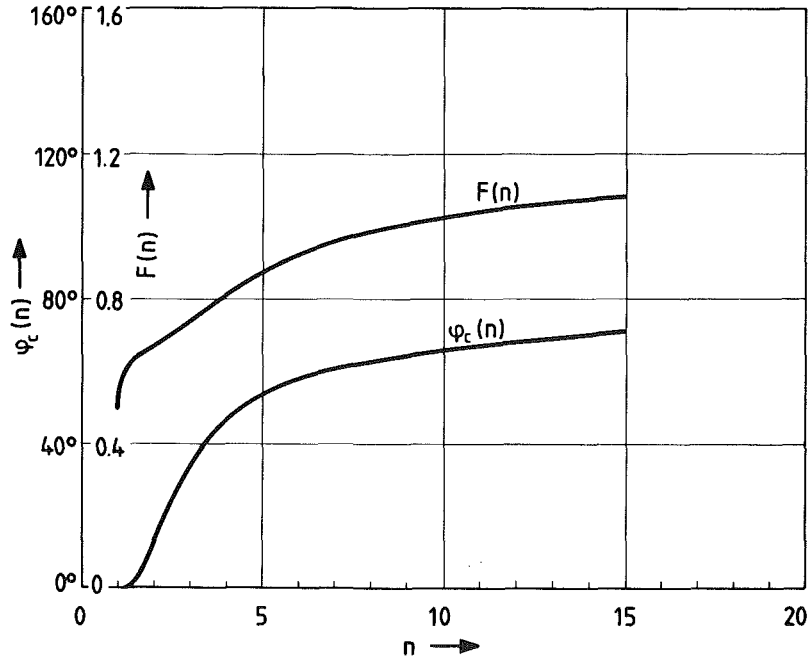


Fig. 5 Evolution of caustics with plastic stress intensity factor K_σ

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

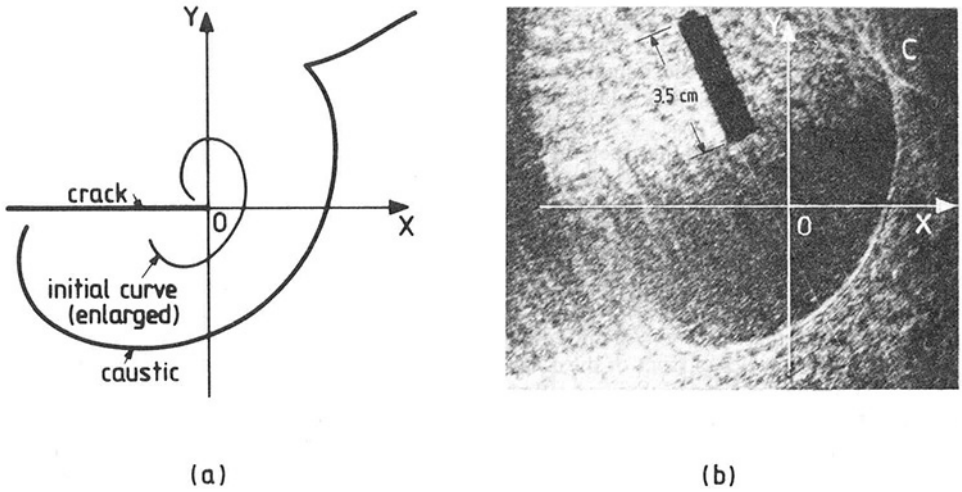


Fig. 6 Theoretically (*a*) and experimentally (*b*) obtained caustics at the tip of a cracked plate of stainless steel under antiplane longitudinal shear.

slope of the thickness outside the plastic zone is insignificant, as compared with the variation of slope in the transition region at the vicinity of the elastic-plastic boundary. This explains why no caustics appear outside the plastic region in the experimentally determined caustics, which are always shorter than the respective theoretical caustics (compare Figs. 6(*a*) and 6(*b*).

However, the characteristic difference between elastic and plastic caustics is the existence of a cusp point for the plastic caustics. All measurements for the evaluation of K_σ are related to the cusp point and this fact gives the assurance that all quantities measured are derived from inside the plastic zone.

By plotting the respective caustics for parametric values of the power-hardening coefficient n, as in Fig. 3(*b*), the influence of this coefficient may be derived. Indeed, the two branches of the caustic, forming the cusp point, are intimately related to the type of strain hardening of the material, depending on the respective value of the power-hardening coefficient. Since it has been shown that the orientation of the cusp point relative to the crack depends only on n, the value of the polar angle φ_c of the radius R_c (Fig. 2) to the cusp point and the crack axis, measured on the respective caustic, yields the value of the power-hardening coefficient of the material.

Evaluation of the Plastic Stress Intensity Factor

In order to evaluate the plastic stress intensity factor K_σ in antiplane shear from the respective caustic, it is sufficient to measure the coordinates of a characteristic point of the caustic. As such point, the cusp point of the caustic is always selected, since this point is indicative of the existence of a plastic zone around the crack tip and, therefore, its corresponding point on the initial curve always lies inside the plastic zone. Furthermore, this point is well defined on the caustic. Indeed, the respective caustic for the elastic case does not present a definite cusp point, since this curve is arrested exactly at this point.

The polar coordinates R and φ of any point of the caustic may be derived from relations (22) and (23) and they are given by (Fig. 2):

$$R = \lambda^{\frac{n}{2n+1}} (2z_0 K_\epsilon)^{\frac{n+1}{2n+1}} [F_x^2(\vartheta,n) + F_y^2(\vartheta,n)]^{1/2} \tag{28}$$

$$\varphi = \tan^{-1} \frac{F_y(\vartheta,n)}{F_x(\vartheta,n)} \tag{29}$$

The position of the cusp point of the caustic is defined by the conditions for a local maximum of the polar angle φ, which are $d\varphi/d\vartheta = 0$ and $d^2\varphi/d\vartheta^2 < 0$. Differentiating relation (29) with respect to ϑ and putting it equal to zero, we define a critical value for the polar angle φ_c, corresponding to the cusp point of the caustic, in terms of the respective polar angle ϑ_c of the initial curve and the respective power-hardening coefficient n. This value φ_c, calculated numerically in the computer, when introduced into relation (28) yields the value R_c for the polar radius of the cusp point as follows:

$$R_c = \lambda^{\frac{n}{2n+1}} (2z_0 K_\epsilon)^{\frac{n+1}{2n+1}} F(n) \tag{30}$$

where:

$$F(n) = [F_x^2(\vartheta_c,n) + F_y^2(\vartheta_c,n)]^{1/2} \tag{31}$$

Relations (30) and (31) indicate that, while the quantity $F(n)$ is a function only of the power-hardening coefficient, the radius R_c for the cusp point is a function of n, as well as of the characteristic dimensions of the optical set-up, and the respective plastic stress intensity factor K_σ.

Relation (30), together with equation (9), yields the plastic stress intensity factor K_σ as follows:

$$K_\sigma = \left\{ \frac{R_c^{\frac{2n+1}{n+1}}}{2z_0 \lambda^{\frac{n}{n+1}} [F(n)]^{\frac{2n+1}{n+1}}} \right\}^{1/n} \tag{32}$$

The function $F(n)$, which depends only on the power-hardening coefficient n of the material, is given in Table 1 for typical values of n and is plotted in Fig. 4.

Relation (32) yields the plastic stress intensity factor, normalized to the yield stress τ_0. Since the dimensions of the normalized plastic stress intensity factor K_σ are a length to a power equal to $1/(n+1)$, a comparison of the amount of plasticity at the crack tip is appropriate only for the radius of the plastic zone. The radius of the plastic zone R_0 and the distance β_0 of its center from the crack tip can be determined from relations (10), (11), and (32) and they are given by:

$$R_0 = \left[\frac{1}{2} \frac{R_c^{\frac{2n+1}{n+1}}}{2z_0 \lambda^{\frac{n}{n+1}} [F(n)]^{\frac{2n+1}{n+1}}} \right]^{\frac{n+1}{n}} \tag{33}$$

and

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\beta_0 = \frac{n-1}{n+1} R_0 \tag{34}$$

Evaluation of the Power-Hardening Coefficient

Relation (30) defines the polar angle ϑ_c of the point of the initial curve, which corresponds to the cusp point. It was already shown that ϑ_c depends only on n. Introducing the value of ϑ_c in equation (29) we obtain the value φ_c, which is given by:

$$\varphi_c = \tan^{-1} \frac{F_y[\vartheta_c(n),n]}{F_x[\vartheta_c(n),n]} \tag{35}$$

Relation (35) indicates that the polar angle φ_c of the cusp point is solely dependent of the power-hardening coefficient n.

Moreover, respective points of caustics, corresponding to various amounts of plasticity developed around the crack tip, all lie along the respective polar radii, thus being independent of the size of each caustic. This independence of the polar angle φ_c of the cusp point of the caustic facilitates the application of the method, since it is sufficient to dispose only the curves of Figs. 3(*b*) and 4, in order to evaluate n.

Furthermore, a simple measurement of the respective polar radius R_c suffices for the evaluation of the corresponding plastic stress intensity factor K_σ. Finally, it is obvious from Fig. 3(*b*) that as n increases, the polar angle φ_c increases, although the shape of the respective caustic is continuously shrinking as n takes higher values, tending toward perfect plasticity.

The Influence of the Optical Set-Up on the Shape and Size of Caustics

We shall now study the influence of the characteristic parameters of the optical set-up, that is, the distance z_0 between the specimen and the reference screen, and the optical magnification ratio λ, on the size and shape of the respective caustic.

The system of equations (30) and (35), which defines the cusp point of each caustic, indicates that for a given value of the power-hardening coefficient n, the locus of the cusp point is a polar radius passing through the crack tip. This behavior is similar to the behavior of the plastic stress intensity factor K_σ as shown in Fig. 5, where a series of consecutive caustics was plotted for different values of K_σ.

Based on relations (20), (21), (24), and (25) the foregoing properties can be generalized for every point (X,Y) of the caustic curve and of its initial curve.

For a constant value of angle ϑ and the power-hardening coefficient n, the geometric locus of the points of the initial curve when either λ, z_0, or K_σ vary, is the polar radius passing through the crack tip.

It can be shown from relations (38) and (39) that the same property is valid for every point of a caustic. This fact is illustrated in Figs. 5(*a*,*b*) by the extreme points of the initial curve and the caustic, which lie on lines passing through the crack tip. This means that an increase of the value of the plastic stress intensity factor K_σ, or of the distance z_0, has the same effect on the caustic as does the magnification factor λ, that is, a magnification of the caustic without any change of its shape (Fig. 5). On the contrary, it may be derived from relation (20), that an increase of λ reduces the size of the respective caustic.

Experimental Evidence

In order to check the validity of the theoretical results, a series of experiments was undertaken. Both stainless steel and bronze plates were selected for the preparation of the specimens. The dimensions of the specimens were 0.09m × 0.20m with a thickness of 0.001m and 0.003m, respectively. A special loading frame, as described in reference [1], was used in order to exclude the influence of mode-*I* and mode-*II* deformations to the cracked plates.

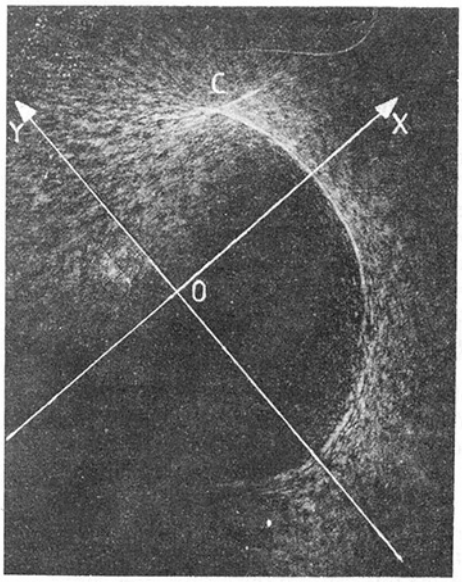


Fig. 7 Experimentally obtained caustic for a bronze plate with $n = 3.5$, under antiplane longitudinal shear.

The optical set-up was designed to have a magnification factor $\lambda = 4$ and a distance $z_0 = 1.50$m. Figure 6(*a*) presents the caustic for the K_{III}-mode of deformation, of a stainless-steel plate, whereas Fig. 7 shows the caustic for a specimen made of bronze.

Both experimental caustics of Figs. 6(*b*) and 7, derived from metallic specimens exhibit a kind of bluntness, these curves being in some parts intermittent and irregular. This lack of sharpness of the caustics is due to a relatively poor polishing of the surface of the specimen and the crystallinity of the metallic substance. However, since the creation of caustics is due to the thickness variation, it is insensitive to the peculiarities of the particular crystallinic structure of the material tested. Thus, because of the size and orientation of crystals and grains, the grain size, and any topical anisotropies only small perturbations may create to the form of caustics. In general, fine-grain metals present a grain size of the order of 10^{-5} to 10^{-4}m, while course-grain sizes are of the order of 10^{-3}m. Therefore, the average grain radii are much below the sizes of the initial curves. For course grains, the caustic will depend on the particular properties of the grain. In this case the problem must be treated according to the theory of caustics for anisotropic materials [10]. However, whether or not the grain boundary intersects the initial curve, the average thickness variation of the specimen, due to loading, always remains the main cause creating the caustic.

The power-hardening coefficient for stainless steel plates was experimentally determined to be $n \simeq 6.0$, whereas, for bronze, this coefficient was found $n \simeq 3.5$. These coefficients were evaluated by simple tension tests. The corresponding theoretically obtained caustics for the stainless steel plates are shown in Fig. 6(*b*). The similarity in shape and size of both respective caustics is evident.

The initial curve indicated in Fig. 6(*a*), which is traced in an enlarged scale, only extends slightly outside the plastic zone and the caustic created by the extremities of the initial curve are missing in the experimental caustics of Figs. 6(*b*) and 7. However, as was already explained, the cusp points of the caustics correspond to points of the initial curves, which

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

always lie inside the plastic zone, since elastic caustics do not present cusp points [1], and their position is sufficient for determining the plastic stress intensity factor.

In order to evaluate the power-hardening coefficient n, the angle φ_c was measured and found $\varphi_c = 57$ deg. This angle corresponds to a value of $n = 5.95$, as can be derived from Table 1. The respective value of n, as determined be a simple tension test, was found to be $n \simeq 6.0$. The discrepancy between these two values is evidently insignificant.

Similarly, for the bronze plates, the angle φ_c was measured to be $\varphi_c = 39.50$ deg, which corresponds to $n = 3.47$, thus very close to its experimental value $n = 3.50$.

For the evaluation of the plastic stress intensity factor in metallic cracked plates under mode-*III*, the experimental caustic of Fig. 6(*a*) for a stainless steel plate was used. The plastic stress intensity factor K_σ was determined by using relation (32) for $\lambda = 4.0$, $z_0 = 1.50$m, and for values of the function $F(n)$ taken from Table 1 for $n = 5.95$. The polar radius R_c at the cusp point of the caustic was measured to be $R_c = 0.0445$m in Fig. 6(*b*). The experimental value of K_σ derived from this test, was equal to $K_\sigma = 0.264\text{m}^{1/6}$.

Introducing this value of the plastic stress intensity factor as well as the values $\lambda = 4.0$, $z_0 = 1.50$m, and $n = 5.95$ in equations (20) for the initial curve, and in equations (22) and (23) for the caustic, the theoretical caustic of Fig. 6(*a*) is plotted, which coincides well with the experimental caustic of Fig. 6(*b*). Then, the approximate values of K_σ and n, experimentally determined from the caustic, may be considered as satisfactory.

Based on the experimentally obtained value of the plastic stress intensity factor $K_\sigma = 0.264\text{m}^{1/6}$ the radius of the plastic zone and the distance β_0 from its center to the crack tip are found to be:

$$R_0 = 0.48 \times 10^{-4}\text{m}$$
$$\beta_0 = 0.35 \times 10^{-4}\text{m}$$

It is worth noting that the maximum value of the radius of the initial curve, evaluated from the experimentally obtained caustic, was found to be $r_{\max} = 0.32 \times 10^{-4}$m. Hence, the radius R_0 of the plastic zone is always larger than the maximum radius $r_{\max}$ of the initial curve in front of the crack tip, wherein lies the point of the initial curve corresponding to the cusp point of the caustic. Since for mode-*III* plastic deformation the center of the plastic zone is always displaced by the quantity β_0 in front of the crack, this explains why the tails of the initial curve, traced in Fig. 6(*a*) by the computer, may lie outside the plastic zone for small-scale yielding where the plastic zone is of equal magnitude with the region of the initial curve.

However, as was already explained, the elastic deformations outside the plastic zone are small, compared to their plastic counterparts, and the lateral deformation in these elastic zones is incapable of creating strain gradients sufficient to form caustics. This explains the fact why the tails of the experimental caustics in Figs. 6(*b*) and 7 are missing, since their respective parts in the initial curves disappear, as lying outside the plastic zone.

Conclusions

In this paper the method of reflected caustics was extended to evaluate the plastic stress intensity factor in metallic cracked plates submitted to antiplane shear for small-scale yielding. It is worthwhile remarking that the caustic for the elastic mode-*III* loading of the plate is always a continuous curve with extremities along the axis of the crack. On the contrary, for an elastic-plastic loading of the plate, the caustic is shifted toward the front of the crack, it is rotated about the crack tip, and always presents a cusp point.

It was further shown that, for increasing values of the power-hardening coefficient, the position of the cusp point is angularly displaced about the crack tip. On the other hand, for increasing loading, the size of the caustic expands uniformly, so that its cusp point always remains along a polar radius. A similar uniform expansion of the caustic takes place when the distance z_0 and magnification ratio λ are increasing.

Finally, it was shown that the polar angle φ_c, of the position vector of the cusp point, directly yields the power-hardening coefficient, whereas the polar radius, together with the polar angle φ_c, define the respective plastic stress intensity factor for any step of external load.

Experiments with metallic cracked plates in antiplane shear yielded values for the plastic stress intensity factor and the power-hardening coefficient which were in good agreement with the theoretical values.

Acknowledgment

The research work contained in this paper was partly supported by the Hercules Cement Company of Greece and by the Hellenic Aluminium Company. The authors are indebted to these institutions for their financial support.

References

1 Theocaris, P. S., "The Reflected Caustics Method for the Evaluation of Mode III Stress-Intensity Factor," *International Journal of Mechanical Sciences*, Vol. 23, 1980, pp. 105–117.

2 Rice, J. R., "Stresses due to a Sharp Notch in a Work-Hardening Elastic-Plastic Material Loaded by Longitudinal Shear," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, 1967, pp. 287–298.

3 Hult, J. A. H., and McClintock, F. A., "Elastic-Plastic Stress and Strain Distribution Around Sharp Notches Under Repeated Shear," *Proceedings Ninth International Congress of Applied Mechanics*, Vol. 8, Brussels, Belgium, 1956, pp. 51–58.

4 Rice, J. R., "Contained Plastic Deformation Near Cracks and Notches Under Longitudinal Shear," *International Journal of Fracture Mechanics*, Vol. 2, June 1966, pp. 426–447.

5 Theocaris, P. S., "The Method of Caustics Applied to Elasticity Problems," in *Optical Methods Applied to Composite Materials, Developments in Composite Materials*, Holister, G. S., Ed., Applied Science Publishers, London, Vol. II, Chapter II, 1981, pp. 27–63.

6 Neuber, H., "Theory of Stress Concentration for Shear-Strain Prismatical Bodies With Arbitrary Nonlinear Stress-Strain Law," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, 1961, pp. 544–550.

7 Hilton, P. D., and Hutchinson, J. W., "Plastic Intensity Factors for Cracked Plates," *Engineering Fracture Mechanics*, Vol. 3, 1970, pp. 435–451.

8 Theocaris, P. S., and Michopoulos, J. G., "A Generalization of Theory of Far-field Caustics by the Catastrophe Theory," *Applied Optics*, Vol. 21, No. 6, 1982, pp. 1080–1091.

9 Theocaris, P. S., "Experimental Determination of the Core Region in Mixed-Mode Fracture," *Proceedings U.S.-Greece Symp. on Mixed-Mode Crack Propagation*, Sih, G., and Theocaris, P. S., Eds., Sigthoff and Noordhoff, 1981, pp. 21–36; also Theocaris, P. S., "The Caustic as a Means to Define the Core Region in Brittle Fracture," *Eng. Fract. Mech.*, Vol. 14, 1981, pp. 353–362.

10 Theocaris, P. S., "Stress Concentrations in Anisotropic Plates by the Method of Caustics," *Journal of Strain Analysis*, Vol. 11, 1979, pp. 154–160.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

L. M. Keer
Mem. ASME

S. Nemat-Nasser
Mem. ASME

A. Oranratnachai

Department of Civil Engineering,
Northwestern University,
The Technological Institute,
Evanston, Ill. 60201

Surface Instability and Splitting in Compressed Brittle Elastic Solids Containing Crack Arrays

Surface instability in brittle solids may occur at relatively small values of the in-plane compressive force if the solid contains shallow cracks parallel to its free surface. The instability may produce surface damage by spallation. Similarly, the buckling load of a longitudinally compressed strip that contains an array of central cracks is affected to a great extent by the size and the relative spacing of these cracks. The instability in this case may result in longitudinal splitting of the strip. To illustrate these phenomena, the compression of an elastic half space and a layer, each containing an array of coplanar equally spaced cracks, is studied for a special class of hypoelastic materials, and the corresponding weakening due to cracks is analytically estimated.

Introduction

The finite deformation due to the horizontal compression of an elastic half space or a layer containing an array of coplanar equally spaced cracks of equal length (Fig. 1) is considered. The objective is to estimate the change in the critical buckling load due to the weakening of the structure by the crack array. It is found that the buckling load decreases dramatically as the distance between the free surface and the crack array diminishes. For shallow cracks, the instability may be interpreted as surface damage due to spallation induced by in-plane compressive forces. Similarly, the buckling load of a longitudinally compressed strip that contains an array of central cracks is affected to a great extent by the size and the relative spacing of these cracks. The instability in this case may result in longitudinal splitting of the strip. Quantitative results are obtained for a class of hypoelastic materials, for which the Jaumann rate of the Kirchhoff stress relates to the deformation rate tensor by an isotropic linear relation similar to Hooke's law. Although other elastoplastic rate-constitutive relations may be considered, the results would be qualitatively the same. Quantitatively, however, different conclusions are expected, as has been illustrated by Dorris and Nemat-Nasser [1] in the case of the buckling of a layer bonded to a half space, and by Horii and Nemat-Nasser [2] in the case of the instability of a half space with frictional materials.

As has been inherent in the pioneering works of Biot (see his book [3] for an account and references to original papers), the formulation of the stability of an initially stressed solid can be most conveniently achieved in terms of incremental quantities referred to the current configuration. Hill [4] has given prominence to the nominal stress and its rate, in this connection. Therefore, in the next section an exact rate formulation is presented in terms of the rate of nominal or the first Piola-Kirchhoff stress, referred to and measured with respect to the current configuration; see, also, Nemat-Nasser [5], and for an account of the theory of small deformations superimposed on large ones, see Green et al. [6]. Then, finite Fourier transforms are used, and the resulting dual series equations are solved by techniques that have been discussed by Snedden and Srivastav [7] and Sneddon and Lowengrub [8]. For the present stability problem this leads to homogeneous integral equations whose eigenvalues are then determined numerically, using a technique that has been employed in a series of papers by Stahl and Keer [9–11].

In the next section the basic equations are presented. Solutions are developed for the whole space with an array of cracks and for a layer with an array of cracks. The formulation for the whole space problem is then used to generate the corresponding equations for the half-space problem containing an array of cracks parallel to the free surface. For all these problems the basic equations are reduced to Fredholm integral equations of the second kind which are then solved numerically.

Formulation

The problems specifically considered are shown in Fig. 1, where a half space and a strip, each containing an array of collinear cracks, are being compressed in the direction of the cracks by a large uniform pressure P. The material is assumed to be hypoelastic, and the deformation is one of plane strain. The theory used in this analysis is based on small deformations superimposed on the large uniform deformation caused by the large uniform pressure (on the same order as the

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July, 1981; final revision, January, 1982. Paper No. 82-WA/APM-16.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

shear modulus); the small deformation will take into account the presence of the crack array. The solutions to Fig. 1 will be obtained by superposition of a whole space solution for an array of cracks in an initially stressed body (Fig. 2(*a*)) with either an undamaged initially stressed half space (Fig. 2(*b*)) or an undamaged initially stressed layer (Fig. 2(*c*)). By appropriate matching of the boundary conditions for these two superposition solutions the surfaces will be free of stress and a stability analysis can be performed.

The stress components employed in this analysis are the first Piola-Kirchhoff stress τ^R_{ab} with $\dot{\tau}^R_{ab}$ being its rate referred to the deformed initially stressed state. In the absence of body forces the equations of equilibrium for the rate problem take on the form (see Biot [3], Hill [4], and Nemat-Nasser [5])

$$\dot{\tau}^R_{ab,a} = 0 \tag{1}$$

where a subscript preceded by a comma denotes partial differentiation with respect to the corresponding coordinate, and where the repeated indices are summed. The incremental stress boundary conditions are

$$\dot{\tau}^R_{ab} n_a = \dot{\tau}_b \tag{2}$$

where $\dot{\tau}_b$ are the components of prescribed traction rates referred to the initial stressed state, and n_a are the components of the outward unit normal of the boundary.

Let $v_a(\mathbf{x},t)$ be the instantaneous velocity components of a typical material point in the current configuration; for the problems considered in this work v_a can be regarded as the incremental deformation. The deformation rate is defined as

$$D_{ab} = \frac{1}{2}(v_{a,b} + v_{b,a}). \tag{3}$$

The stress rate $\dot{\tau}^R_{ab}$ is related to the deformation rate as follows (see Hill [4] and Nemat-Nasser [5]):

$$\dot{\tau}^R_{ab} = \frac{\mathfrak{D}\tau_{ab}}{\mathfrak{D}t} - D_{ac}\sigma_{cb} - D_{bc}\sigma_{ca} + v_{b,c}\sigma_{ca} \tag{4}$$

where $\mathfrak{D}\tau_{ab}/\mathfrak{D}t$ denotes the Jaumann rate of the Kirchhoff stress,[1] and σ_{ab} are the components of the Cauchy stress. For the special hypoelastic solid cosidered here the Jaumann rate of the Kirchhoff stress is assumed to be related isotropically and linearly to the deformation rate

$$\frac{\mathfrak{D}\tau_{ab}}{\mathfrak{D}t} = 2\mu D_{ab} + \lambda D_{cc}\delta_{ab} \tag{5}$$

where λ and μ are the Lamé constants for liner elasticity, and δ_{ab} is Kronecker's delta. By treating σ_{ab} in equation (4) as an initial stress, the basic equations for the plane problems we wish to consider are

$$\dot{\tau}^R_{11} = (2\mu+\lambda+P)v_{1,1} + \lambda v_{2,2} \tag{6}$$

$$\dot{\tau}^R_{22} = (2\mu+\lambda)v_{2,2} + \lambda v_{1,1} \tag{7}$$

$$\dot{\tau}^R_{12} = (\mu+P/2)v_{1,2} + (\mu-P/2)v_{2,1} \tag{8}$$

$$\dot{\tau}^R_{21} = (\mu+P/2)(v_{1,2} + v_{2,1}). \tag{9}$$

As was shown by Dorris and Nemat-Nasser [1], if the displacements have the following representation

$$v_1 = i\omega[Ae^{-\gamma_1\omega y} + Be^{-\gamma_2\omega y} + Ce^{\gamma_1\omega y} + De^{\gamma_2\omega y}]e^{i\omega x} + R_1 \tag{10}$$

$$v_2 = \omega[-k_1\gamma_1 Ae^{-\gamma_1\omega y} - k_2\gamma_2 Be^{-\gamma_2\omega y} + k_1\gamma_1 Ce^{\gamma_1\omega y} + k_2\gamma_2 De^{\gamma_2\omega y}]e^{i\omega x} + R_2 \tag{11}$$

where

$$k_1 = 1, \quad k_2 = (1+\alpha)/(1-\alpha) \tag{12}$$

$$\gamma_1 = 1, \quad \gamma_2 = [(2+\eta+2\alpha)(1-\alpha)/(2+\eta)(1+\alpha)]^{1/2} \tag{13}$$

and

$$\alpha = P/2\mu, \quad \eta = \lambda/\mu \tag{14}$$

then the equations of equilibrium, equation (1), will be automatically satisfied. Here, A, B, C, D, R_1, R_2 are constants to be determined from the boundary conditions, and ω represents the wavelength of the buckled shape. The solution for the basic problems as shown in Fig. 2 will be constructed next.

Solution for the Whole Space With an Array of Cracks

We consider two half spaces, $y \lessgtr 0$, where the half space $y < 0$ will be labeled as "1" and the half space $y > 0$ will be labeled as "2." If in equations (10) and (11) we set $\omega = n\pi/L$ and use the principle of superposition, then solutions for the

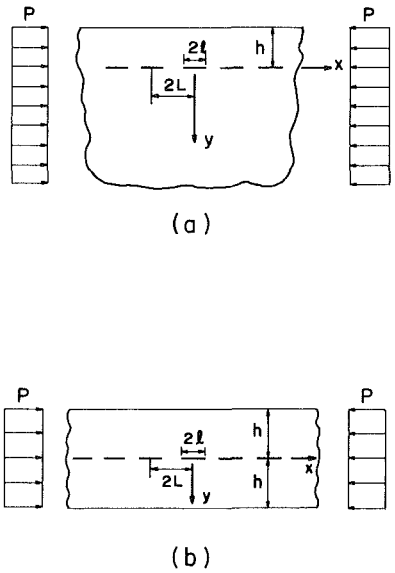


Fig. 1 (a) A half space with an array of collinear cracks subjected to a large horizontal compression *P*; and (*b*) a strip with an array of collinear cracks along the center line subjected to a large horizontal compression *P*

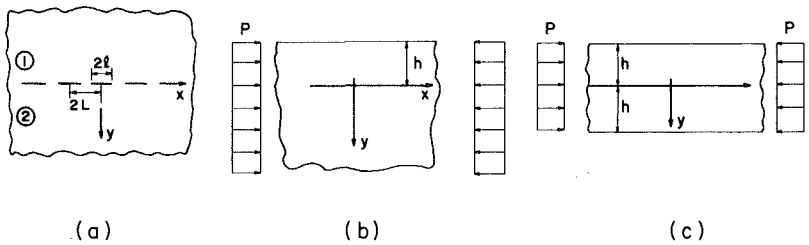


Fig. 2 (*a*) Geometry of an array of collinear cracks in an infinite plane; (*b*) geometry of a half space subjected to the horizontal compression *P*; and (c) geometry of a strip subjected to the horizontal compression *P*

[1] The Kirchhoff stress relates to the Cauchy stress by $\tau_{ab} = \rho_0/\rho\,\sigma_{ab}$, where ρ_0 is the reference, and ρ is the current mass density.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

upper and lower half spaces can be obtained that satisfy the following continuity conditions at $y = 0$:

$$\dot{\tau}_{22}^{R(1)} = \dot{\tau}_{22}^{R(2)} \quad -\infty < x < \infty \tag{15}$$

$$\dot{\tau}_{21}^{R(1)} = \dot{\tau}_{21}^{R(2)} \quad -\infty < x < \infty. \tag{16}$$

These solutions are as follows:

Half space $y \le 0$:

$$v_1^{(1)} = \sum_{n=1}^{\infty} \frac{in\pi}{2L} \left[\tilde{a}_n \left(e^{\gamma_1 n\pi y/L} - \frac{K_1}{K_2} e^{\gamma_2 n\pi y/L} \right) + \tilde{b}_n \left(e^{\gamma_1 n\pi y/L} - \frac{W_1}{W_2} e^{\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} + R_1 \tag{17}$$

$$v_2^{(1)} = \sum_{n=1}^{\infty} \frac{n\pi}{2L} \left[\tilde{a}_n \left(k_1\gamma_1 e^{\gamma_1 n\pi y/L} - \frac{K_1}{K_2} k_2\gamma_2 e^{\gamma_2 n\pi y/L} \right) + \tilde{b}_n \left(k_1\gamma_1 e^{\gamma_1 n\pi y/L} = \frac{W_1}{W_2} k_2\gamma_2 e^{\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} + R_2 \tag{18}$$

$$\dot{\tau}_{22}^{R(1)}/\mu = \frac{1}{2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 \left[\tilde{a}_n K_1 \left(e^{\gamma_1 n\pi y/L} - e^{\gamma_2 n\pi y/L} \right) + \frac{\tilde{b}_n}{W_2} \left(K_1 W_2 e^{\gamma_1 n\pi y/L} - K_2 W_1 e^{\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} \tag{19}$$

$$\dot{\tau}_{21}^{R(1)}/\mu = (1+\alpha) \sum_{n=1}^{\infty} \frac{i}{2} \left(\frac{n\pi}{L} \right)^2 \left[\frac{\tilde{a}_n}{K_2} \left(K_2 W_1 e^{\gamma_1 n\pi y/L} - K_1 W_2 e^{\gamma_2 n\pi y/L} \right) + \tilde{b}_n W_1 \left(e^{\gamma_1 n\pi y/L} - e^{\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L}. \tag{20}$$

Half space $y \ge 0$:

$$v_1^{(2)} = \sum_{n=1}^{\infty} \frac{in\pi}{2L} \left[\tilde{a}_n \left(-e^{-\gamma_1 n\pi y/L} + \frac{K_1}{K_2} e^{-\gamma_2 n\pi y/L} \right) + \tilde{b}_n \left(e^{-\gamma_1 n\pi y/L} - \frac{W_1}{W_2} e^{-\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} \tag{21}$$

$$v_2^{(2)} = \sum_{n=1}^{\infty} \frac{n\pi}{2L} \left[\tilde{a}_n \left(k_1\gamma_1 e^{-\gamma_1 n\pi y/L} - \frac{K_1}{K_2} k_2\gamma_2 e^{-\gamma_2 n\pi y/L} \right) + \tilde{b}_n \left(-k_1\gamma_1 e^{-\gamma_1 n\pi y/L} + \frac{W_1}{W_2} k_2\gamma_2 e^{-\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} \tag{22}$$

$$\dot{\tau}_{22}^{R(2)}/\mu = \frac{1}{2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 \left[\tilde{a}_n K_1 \left(-e^{-\gamma_1 n\pi y/L} + e^{-\gamma_2 n\pi y/L} \right) + \frac{\tilde{b}_n}{W_2} \left(K_1 W_2 e^{-\gamma_1 n\pi y/L} - K_2 W_1 e^{-\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L} \tag{23}$$

$$\dot{\tau}_{21}^{R(2)}/\mu(1+\alpha) \sum_{n=1}^{\infty} \frac{i}{2} \left(\frac{n\pi}{L} \right)^2 \left[\frac{\tilde{a}_n}{K_2} \left(K_2 W_1 e^{-\gamma_1 n\pi y/L} - K_1 W_2 e^{-\gamma_2 n\pi y/L} \right) + \tilde{b}_n W_1 \left(-e^{-\gamma_1 n\pi y/L} + e^{-\gamma_2 n\pi y/L} \right) \right] e^{in\pi x/L}. \tag{24}$$

In equations (17)–(24) we have

$$K_1 = (2+\eta) k_1 \gamma_1^2 - \eta, \quad K_2 = (2+\eta) k_2 \gamma_2^2 - \eta \tag{25}$$

$$W_1 = \gamma_1 (1+k_1), \quad W_2 = \gamma_2 (1+k_2). \tag{26}$$

The boundary conditions that correspond to the crack array can be written in the following form:

$$\dot{\tau}_{22}^{R}/\mu = f(x) \quad 0 < |x - 2nL| < l \tag{27}$$

$$\dot{\tau}_{21}^{R}/\mu = g(x) \quad 0 < |x - 2nL| < l \tag{28}$$

$$v_1^{(1)} = v_1^{(2)}, \; v_2^{(1)} = v_2^{(2)} \quad l < |x - 2nL| < L \tag{29}$$

which express the fact that each crack in the array is loaded by normal and shear stress rates, and that the velocity field is continuous outside of the cracks.

If we write $f(x) = f_e(x) + if_o(x)$ and $g(x) = g_e(x) + ig_o(x)$, where e and o refer to the symmetric and antisymmetric parts of the traction acting on the crack faces, then the boundary value problem, equations (27)–(29), can be written as the following two sets of mixed boundary value problems:

(a) Symmetric Part.

$$\frac{\mathfrak{L}}{2W_2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 \tilde{b}_n \cos(n\pi x/L) = f_e(x) \quad 0 < x < l \tag{30}$$

$$\tilde{b}_0 + \frac{\gamma_1\gamma_2(k_1 - k_2)}{W_2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \tilde{b}_n \cos(n\pi x/L) = 0 \quad l < x < L \tag{31}$$

$$\frac{\mathfrak{L}(1+\alpha)}{2K_2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 \tilde{a}_n \sin(n\pi x/L) = g_e(x) \quad 0 < x < l \tag{32}$$

$$\frac{(K_1 - K_2)}{K_2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \tilde{a}_n \sin(n\pi x/L) = 0 \quad l < x < L. \tag{33}$$

(b) Antisymmetric Part.

$$\frac{\mathfrak{L}}{2W_2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 b_n' \sin(n\pi x/L) = f_o(x) \quad 0 < x < l \tag{34}$$

$$\gamma_1\gamma_2(k_1 - k_2) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \frac{b_n'}{W_2} \sin(n\pi x/L) = 0 \quad l < x < L \tag{35}$$

$$-\frac{\mathfrak{L}}{2}(1+\alpha) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right)^2 \frac{a_n'}{K_2} \cos(n\pi x/L) = g_o(x) \quad 0 < x < l \tag{36}$$

$$-(K_1 - K_2) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \frac{a_n'}{K_2} \cos(n\pi x/L) + a_0' = 0 \quad l < x < L \tag{37}$$

where

$$\mathfrak{L} = K_1 W_2 - K_2 W_1. \tag{38}$$

Solution for the Symmetric Part (a). Let $\tilde{x} = \pi x/L$ and $\tilde{l} = \pi l/L$, and use the method employed by Sneddon and Srivastav [7] (see also Chapter 2 of [8]). We let

$$\tilde{b}_0 + \gamma_1\gamma_2(k_1 - k_2) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \frac{\tilde{b}_n}{W_2} \cos(n\tilde{x}) = \begin{cases} \cos(\tilde{x}/2) \displaystyle\int_{\tilde{x}}^{\tilde{l}} \frac{\phi_1(t)\,dt}{[\cos\tilde{x} - \cos t]^{1/2}} & 0 < \tilde{x} < \tilde{l}, \\ 0 & \tilde{l} < \tilde{x} < \pi, \end{cases} \tag{39}$$

$$(K_1 - K_2) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L} \right) \frac{\tilde{a}_n}{K_2} \sin(n\tilde{x}) = \begin{cases} \sin(\tilde{x}/2) \displaystyle\int_{\tilde{x}}^{\tilde{l}} \frac{\phi_2(t)\,dt}{[\cos\tilde{x} - \cos t]^{1/2}} & 0 < \tilde{x} < \tilde{l}, \\ 0 & \tilde{l} < \tilde{x} < \pi. \end{cases} \tag{40}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Using these definitions, the Fourier coefficients $\tilde{a}_n$, $\tilde{b}_n$, $\tilde{b}_0$ are given in terms of the Legendre polynomial as follows:

$$\tilde{b}_0 = \frac{1}{\sqrt{2}} \int_0^{\tilde{l}} \phi_1(t)\, dt \tag{41}$$

$$\tilde{b}_n = \frac{W_2 L}{n\pi\sqrt{2}\gamma_1\gamma_2(k_1-k_2)} \int_0^{\tilde{l}} \phi_1(t)\,[P_{n-1}(\cos t) + P_n(\cos t)]\,dt \tag{42}$$

$$\tilde{a}_n = \frac{K_2 L}{n\pi\sqrt{2}(K_1-K_2)} \int_0^{\tilde{l}} \phi_2(t)\,[P_{n-1}(\cos t) - P_n(\cos t)]\,dt. \tag{43}$$

Substituting equations (42) and (43) into (30) and (32) and use of the identities

$$\frac{1}{\sqrt{2}} \sum_{n=1}^{\infty} [P_n(\cos t) + P_{n-1}(\cos t)]\sin(n\tilde{x}) = \frac{\cos(\tilde{x}/2)H(\tilde{x}-t)}{[\cos t - \cos\tilde{x}]^{1/2}} \tag{44}$$

$$\frac{1}{\sqrt{2}} \sum_{n=1}^{\infty} [P_{n-1}(\cos t) - P_n(\cos t)]\cos(n\tilde{x}) = \frac{1}{\sqrt{2}} - \frac{\sin(\tilde{x}/2)H(\tilde{x}-t)}{[\cos t - \cos\tilde{x}]^{1/2}} \tag{45}$$

(where $H(x)$ is the Heaviside function) lead to the following equations:

$$\frac{\mathcal{L}}{2\gamma_1\gamma_2(k_1-k_2)} \int_0^{\tilde{x}} \frac{\phi_1(t)\cos(\tilde{x}/2)}{[\cos t - \cos\tilde{x}]^{1/2}}\, dt = \int_0^{\tilde{x}} f_e(s)\, ds = F_e(\tilde{x}) \quad 0<\tilde{x}<\tilde{l} \tag{46}$$

$$\frac{\mathcal{L}(1+\alpha)}{2(K_1-K_2)} \left\{ \frac{1}{\sqrt{2}} \int_0^{\tilde{l}} \phi_2(t)\, dt - \int_0^{\tilde{x}} \frac{\phi_2(t)\sin(\tilde{x}/2)}{[\cos t - \cos\tilde{x}]^{1/2}}\, dt \right\} = \int_0^{\tilde{x}} g_e(s)\, ds = G_e(\tilde{x}) \quad 0<\tilde{x}<\tilde{l}. \tag{47}$$

The solution of integral equations (46) and (47) yields

$$\frac{\mathcal{L}\pi}{4\gamma_1\gamma_2(k_1-k_2)}\phi_1(z) = \frac{d}{dz}\int_0^z \frac{\sin(\tilde{x}/2)F_e(\tilde{x})}{[\cos\tilde{x}-\cos z]^{1/2}}\, d\tilde{x} \quad 0<z<\tilde{l} \tag{48}$$

$$\frac{\mathcal{L}\pi(1+\alpha)}{4(K_1-K_2)}\phi_2(z) = \frac{d}{dz}\int_0^z \frac{\cos(\tilde{x}/2)G_e(\tilde{x})}{[\cos\tilde{x}-\cos z]^{1/2}}\, d\tilde{x} \quad 0<z<\tilde{l}. \tag{49}$$

Solution for the Antisymmetric Part (*b*). For this case we let

$$\gamma_1\gamma_2(k_1-k_2)\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)\frac{b_n'}{W_2}\sin(n\tilde{x}) = \begin{cases} \sin(\tilde{x}/2)\displaystyle\int_{\tilde{x}}^{\tilde{l}} \frac{\psi_1(t)\,dt}{[\cos\tilde{x}-\cos t]^{1/2}} & 0<\tilde{x}<\tilde{l}, \\ 0 & \tilde{l}<\tilde{x}<\pi, \end{cases} \tag{50}$$

$$a_0' - (K_1-K_2)\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)\frac{a_n'}{K_2}\cos(n\tilde{x}) = \begin{cases} \cos(\tilde{x}/2)\displaystyle\int_{\tilde{x}}^{\tilde{l}} \frac{\psi_2(t)\,dt}{[\cos\tilde{x}-\cos t]^{1/2}} & 0<\tilde{x}<\tilde{l}, \\ 0 & \tilde{l}<\tilde{x}<\pi. \end{cases} \tag{51}$$

The Fourier coefficients in this case, a_n', b_n', are given in the form

$$b_n' = \frac{W_2 L}{n\pi\sqrt{2}\gamma_1\gamma_2(k_1-k_2)} \int_0^{\tilde{l}} \psi_1(t)\,[P_{n-1}(\cos t) - P_n(\cos t)]\,dt \tag{52}$$

$$a_n' = \frac{-K_2 L}{n\pi\sqrt{2}(K_1-K_2)} \int_0^{\tilde{l}} \psi_2(t)\,[P_{n-1}(\cos t) + P_n(\cos t)]\,dt \tag{53}$$

$$a_0' = \frac{1}{\sqrt{2}} \int_0^{\tilde{l}} \psi_2(t)\, dt. \tag{54}$$

After some manipulations similar to those done earlier for the symmetrical part, we find that

$$\frac{\mathcal{L}\pi}{4\gamma_1\gamma_2(k_1-k_2)}\psi_1(z) = \frac{d}{dz}\int_0^z \frac{F_o(\tilde{x})\cos(\tilde{x}/2)d\tilde{x}}{[\cos\tilde{x}-\cos z]^{1/2}} \tag{55}$$

$$\frac{\mathcal{L}\pi(1+\alpha)}{4(K_1-K_2)}\psi_2(z) = \frac{d}{dz}\int_0^z \frac{G_o(\tilde{x})\sin(\tilde{x}/2)d\tilde{x}}{[\cos\tilde{x}-\cos z]^{1/2}} \tag{56}$$

where

$$F_o(\tilde{x}) = \int_0^x f_o(s)\, ds, \quad G_o(\tilde{x}) = \int_0^x g_o(s)\, ds. \tag{57}$$

Solution for a Half Space With an Array of Cracks

For a half space with coordinate axes as shown in Fig. 2(*b*) ($y + h \geq 0$) it can be shown that the incremental stress components can be written as

$$\dot{\tau}_{22}^{R\prime}/\mu = \sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2 \left[K_1 A_n e^{-\gamma_1 n\pi(y+h)/L} + K_2 B_n e^{-\gamma_2 n\pi(y+h)/L}\right] e^{in\pi x/L} \tag{58}$$

$$\dot{\tau}_{21}^{R\prime}/\mu = -i(1+\alpha)\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2 \left[W_1 A_n e^{-\gamma_1 n\pi(y+h)/L} + W_2 B_n e^{-\gamma_2 n\pi(y+h)/L}\right] e^{in\pi x/L} \tag{59}$$

where equations (58) and (59) are derived from a suitably modified form of (10) and (11) by replacing y by $y + h$ and retaining only negative exponentials.

The solution for the problem as shown in Fig. 1(*a*) can now be obtained by the superposition of solutions for the infinite plane with the crack array and for the undamaged half space provided that the following matching conditions are satisfied:

On $y = -h$:

$$\dot{\tau}_{22}^{R\prime}(x,-h) + \dot{\tau}_{22}^{R(1)}(x,-h) = 0 \quad -\infty<x<\infty \tag{60}$$

$$\dot{\tau}_{21}^{R\prime}(x,-h) + \dot{\tau}_{21}^{R(1)}(x,-h) = 0 \quad -\infty<x<\infty \tag{61}$$

On $y = 0$:

$$\dot{\tau}_{22}^{R\prime}(x,0) + \dot{\tau}_{22}^{R(1)}(x,0) = 0 \quad 0<|x-2nL|<l \tag{62}$$

$$\dot{\tau}_{21}^{R\prime}(x,0) + \dot{\tau}_{21}^{R(1)}(x,0) = 0 \quad 0<|x-2nL|<l. \tag{63}$$

For this case we consider only the possibility of a symmetric solution. Equations (60) and (61) lead to solutions of A_n and B_n in terms of $\tilde{a}_n$ and $\tilde{b}_n$ as

$$A_n = \frac{1}{2\mathcal{L}}\left\{\tilde{a}_n\left[2K_1W_2E_n' - E_n(K_1W_2+K_2W_1)\right] + \tilde{b}_n\left[2K_2W_1E_n' - E_n(K_1W_2+K_2W_1)\right]\right\} \tag{64}$$

$$B_n = \frac{1}{2\mathcal{L}}\left\{\tilde{a}_n\left[2K_1W_1E_n - \frac{K_1E_n'}{K_2}(K_1W_2+K_2W_1)\right] + \tilde{b}_n\left[2K_1W_1E_n - \frac{W_1E_n'}{W_2}(K_1W_2+K_2W_1)\right]\right\} \tag{65}$$

where

$$E_n = e^{-\gamma_1 n\pi h/L}, \quad E_n' = e^{-\gamma_2 n\pi h/L}. \tag{66}$$

Then from the use of equations (30) and (32), and (58) and (59), equations (62) and (63) become

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$f_e(x)=\frac{1}{2\mathfrak{L}}\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2\left\{\tilde{a}_n K_1(K_1W_2+K_2W_1)(E_n-E_n')^2+\frac{\tilde{b}_n}{W_2}\left[(K_1W_2E_n-K_2W_1E_n')^2+K_1K_2W_1W_2(E_n-E_n')^2\right]\right\}\cos(n\pi x/L), \tag{67}$$

$$g_e(x)=\frac{(1+\alpha)}{2\mathfrak{L}}\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2\left\{\frac{\tilde{a}_n}{K_2}\left[(K_2W_1E_n-K_1W_2E_n')^2+K_1K_2W_1W_2(E_n-E_n')^2\right]+\tilde{b}_nW_1(K_1W_2+K_2W_1)(E_n-E_n')^2\right\}\sin(n\pi x/L). \tag{68}$$

Using the same technique that led to equations (48) and (49) now leads to the following coupled pairs of homogeneous Fredholm integral equations of the second kind:

$$\frac{2\mathfrak{L}}{\gamma_1\gamma_2(k_1-k_2)}\phi_1(z)-\frac{1}{\mathfrak{L}\gamma_1\gamma_2(k_1-k_2)}\int_0^{\bar{l}}\phi_1(t)M_{11}(t,z)\,dt-\frac{K_1K_2(K_1W_2+K_2W_1)}{\mathfrak{L}(K_1-K_2)}\int_0^{\bar{l}}\phi_2(t)M_{12}(t,z)\,dt=0\quad 0<z<\bar{l} \tag{69}$$

$$\frac{2\mathfrak{L}}{K_1-K_2}\phi_2(z)-\frac{1}{\mathfrak{L}(K_1-K_2)}\int_0^{\bar{l}}\phi_2(t)M_{22}(t,z)\,dt-\frac{W_1W_2(K_1W_2+K_2W_1}{\mathfrak{L}\gamma_1\gamma_2(k_1-k_2)}\int_0^{\bar{l}}\phi_1(t)M_{21}(t,z)\,dt=0\quad 0<z<\bar{l} \tag{70}$$

where

$$M_{11}(t,z)=\left(\frac{1-\cos z}{1+\cos z}\right)^{1/2}\sum_{n=1}^{\infty}\{n[K_1W_2E_n-K_2W_1E_n')^2+K_1K_2W_1W_2(E_n-E_n')^2]\times[P_{n-1}(\cos t)+P_n(\cos t)][P_{n-1}(\cos z)+P_n(\cos z)]\} \tag{71}$$

$$M_{12}(t,z)=\left(\frac{1-\cos z}{1+\cos z}\right)^{1/2}\times\sum_{n=1}^{\infty}\{n(E_n-E_n')^2[P_{n-1}(\cos t)-P_n(\cos t)]\times[P_{n-1}(\cos z)+P_n(\cos z)]\} \tag{72}$$

$$M_{21}(t,z)=\left(\frac{1+\cos z}{1-\cos z}\right)^{1/2}\sum_{n=1}^{\infty}\{n(E_n-E_n')^2[P_{n-1}(\cos t)+P_n(\cos t)]\times[P_{n-1}(\cos z)-P_n(\cos z)]\} \tag{73}$$

$$M_{22}(t,z)=\left(\frac{1+\cos z}{1-\cos z}\right)^{1/2}\sum_{n=1}^{\infty}\{n[(K_2W_1E_n-K_1W_2E_n')^2+K_1K_2W_1W_2(E_n-E_n')^2]\times[P_{n-1}(\cos t)-P_n(\cos t)][P_{n-1}(\cos z)-P_n(\cos z)]\}. \tag{74}$$

Solution for a Layer

For a layer (see Fig. 1(*b*)) it is appropriate to consider two specific forms of buckled shapes: (*a*) symmetric shape in which the lateral displacements are antisymmetric with respect to the center line of the layer and (*b*) antisymmetric shape in which displacements parallel to the layer center line are antisymmetric with respect to it (sideways). The boundary conditions appropriate to these respective symmetries are as follows:

(*a*) **Symmetric Problem ($y = 0$).**

$$\dot{\tau}_{21}^R=0\quad -\infty<x<\infty \tag{75}$$

$$\dot{\tau}_{22}^R=0\quad 0<|x-2nL|<l \tag{76}$$

$$v_2=0\quad l<|x-2nL|<L. \tag{77}$$

(*b*) **Antisymmetric Problem ($y = 0$).**

$$\dot{\tau}_{22}^R=0\quad -\infty<x<\infty \tag{78}$$

$$\dot{\tau}_{21}^R=0\quad 0<|x-2nL|<l \tag{79}$$

$$v_1=0\quad l<|x-2nL|<L. \tag{80}$$

Symmetric Case (*a*). The stress rates in the undamaged layer, as shown in Fig. 2(*c*) that are suitable for the boundary conditions (75)–(77) are

$$\dot{\tau}_{22}^R/\mu=\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2\left\{A_nK_1\cosh(\gamma_1n\pi y/L)+B_nK_2\cosh(\gamma_2n\pi y/L)\right\}\cos(n\pi x/L) \tag{81}$$

$$\dot{\tau}_{21}^R/\mu=(1+\alpha)\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2\left\{-A_nW_1\sinh(\gamma_1n\pi y/L)-B_nW_2\sinh(\gamma_2n\pi y/L)\right\}\sin(n\pi x/L). \tag{82}$$

To satisfy equation (75), it is seen from equation (20) that $\tilde{a}_n = 0$, which implies that $\phi_2 = 0$. Upon applying boundary conditions similar to those of (60) and (61), we obtain the expressions for A_n and B_n in terms of $\tilde{b}_n$ as follows:

$$A_n=\frac{2\tilde{b}_n}{\Delta}[\sinh(\gamma_2n\pi h/L)(-K_1W_2E_n+K_2W_1E_n')-K_2W_1\cosh(\gamma_2n\pi h/L)(E_n-E_n')] \tag{83}$$

$$B_n=\frac{2\tilde{b}_n}{\Delta}[K_1W_1\cosh(\gamma_1n\pi h/L)(E_n-E_n')-\frac{W_1}{W_2}\sinh(\gamma_1n\pi h/L)(-K_1W_2E_n+K_2W_1E_n')] \tag{84}$$

where

$$\Delta=[\mathfrak{L}(1-E_n^2E_n'^2)+(K_1W_2+K_2W_1)(E_n^2-E_n'^2)]/(E_nE_n'). \tag{85}$$

Using equations analogous to equation (62), we obtain

$$f_e(x)=-\frac{2}{\Delta}\sum_{n=1}^{\infty}\left(\frac{n\pi}{L}\right)^2\frac{\tilde{b}_n}{W_2}\left\{(-K_1W_2E_n+K_2W_1E_n')[K_1W_2\sinh(\gamma_2n\pi h/L)-K_2W_1\sinh(\gamma_1n\pi h/L)]+K_1K_2W_1W_1(E_n-E_n')[\cosh(\gamma_1n\pi h/L)-\cosh(\gamma_2n\pi h/L)\right\}\cos(n\pi x/L). \tag{86}$$

Substituting the definition of $\tilde{b}_n$, equation (42), into the foregoing and subsequently into (48), yields

$$\mathfrak{L}\phi_1(z)+\int_0^{\bar{l}}\phi_1(t)Q_1(,z)\,dt=0\quad 0<z<\bar{l} \tag{87}$$

where

$$Q_1(t,z)=\sqrt{\frac{1-\cos z}{1+\cos z}}\sum_{n=1}^{\infty}\frac{2n}{\Delta}\left\{(-K_1W_2E_n+K_2W_1E_n')\times[K_1W_2\sinh(\gamma_2n\pi h/L)-K_2W_1\sinh(\gamma n\pi h/L)]+K_1K_2W_1W_2(E_n-E_n')[\cosh(\gamma_1n\pi h/L)-\cosh(\gamma_2n\pi h/L)]\right\}\times\left\{[P_{n-1}(\cos t)+P_n(\cos t)][P_{n-1}(\cos z)+P_n(\cos z)]\right\}. \tag{88}$$

Antisymmetric Case (*b*). The incremental stress components in this case are

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\tau_{22}^R/\mu = \sum_{n=1}^{\infty} \left(\frac{n\pi}{L}\right)^2 \left[C_n K_1 \sinh(\gamma_1 n\pi y/h) + D_n K_2 \sinh(\gamma_2 n\pi y/h)\right] \sin(n\pi x/L) \quad (89)$$

$$\tau_{21}^R/\mu = (1+\alpha) \sum_{n=1}^{\infty} \left(\frac{n\pi}{L}\right)^2 \left[C_n W_1 \cosh(\gamma_1 n\pi y/h) + D_n W_2 \cosh(\gamma_2 n\pi y/h)\right] \cos(n\pi x/L) \quad (90)$$

where the constants C_n, D_n are in terms of a_n' as follows:

$$C_n = \frac{2a_n'}{\Delta}\left[K_1 W_2 \cosh(\gamma_2 n\pi h/L)\,(E_n - E_n') - \sinh(\gamma_2 n\pi h/L)\,(K_1 W_2 E_n' - K_2 W_1 E_n)\right] \quad (91)$$

$$D_n = \frac{2a_n'}{\Delta}\left[\frac{K_1}{K_2}\sinh(\gamma_1 n\pi h/L)\,(K_1 W_2 E_n' - K_2 W_1 E_n) - K_1 W_1 \cosh(\gamma_1 n\pi h/L)\,(E_n - E_n')\right]; \quad (92)$$

b_n' is zero in this case in order to satisfy equation (78). Using in this case an equation analogous to (63) and the foregoing lead to

$$g_o(x) = \frac{-2(1+\alpha)}{\Delta} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L}\right)^2 \frac{a_n'}{K_2}\left\{(K_1 W_2 E_n' - K_2 W_1 E_n) \times [K_1 W_2 \sinh(\gamma_1 n\pi h/L) - K_2 W_1 \sinh(\gamma_2 n\pi h/L)] + K_1 K_2 W_1 W_2 (E_n - E_n')[\cosh(\gamma_2 n\pi h/L) - \cosh(\gamma_1 n\pi h/L)]\right\} \cos(n\pi x/L) \quad (93)$$

and substituting this result together with (53) into equation (56) leads to

$$\mathcal{L}\psi_2(z) + \int_0^{\bar{l}} \psi_2(t)\, Q_2(t,z)\, dt = 0 \quad 0 < z < \bar{l} \quad (94)$$

where

$$Q_2(t,z) = \sqrt{\frac{1-\cos z}{1+\cos z}} \sum_{n=1}^{\infty} \frac{2n}{\Delta}\left\{(-K_1 W_2 E_n' + K_2 W_1 E_n) \times [K_1 W_2 \sinh(\gamma_1 n\pi h/L) - K_2 W_1 \sinh(\gamma_2 n\pi h/L)] + K_1 K_2 W_1 W_2 (E_n - E_n')[\cosh(\gamma_1 n\pi h/L) - \cosh(\gamma_2 n\pi h/L)]\right\} \times \left\{[P_{n-1}(\cos t) + P_n(\cos t)][P_{n-1}(\cos z) + P_n(\cos z)]\right\}. \quad (95)$$

Results and Conclusions

The integral equations are solved by using a quadrature scheme to express the integrals in discrete form. Since the integral equations are of the second kind the collocation points and the points used in the integration scheme are taken to be the same. A system of n homogeneous linear algebraic equations for the n unknowns is developed. If the unknowns are to represent a nontrivial solution, then the matrix formed by the coefficients must have a determinant equal to zero. This is achieved by fixing the elastic constants and the geometry and varying the load parameter $\alpha = P/2\mu$ until the determinant becomes zero [9–11].

For the present analysis the compression of a half space containing an array is considered first. In Fig. 3 the buckling load α is shown plotted as a function of the depth, h, of the cracks from the free surface and the crack length, l; we note

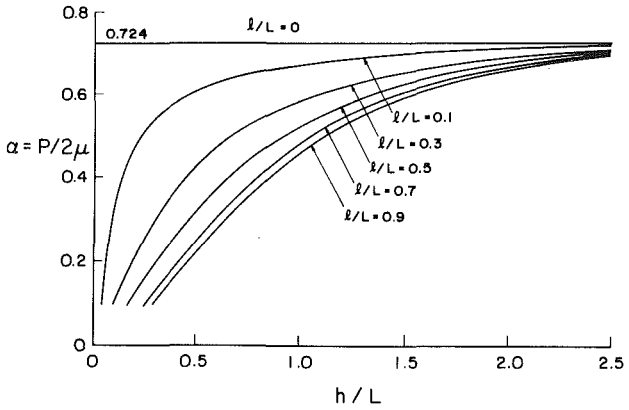


Fig. 3 Buckling load of a half space containing an array of collinear cracks (ν = 0.3)

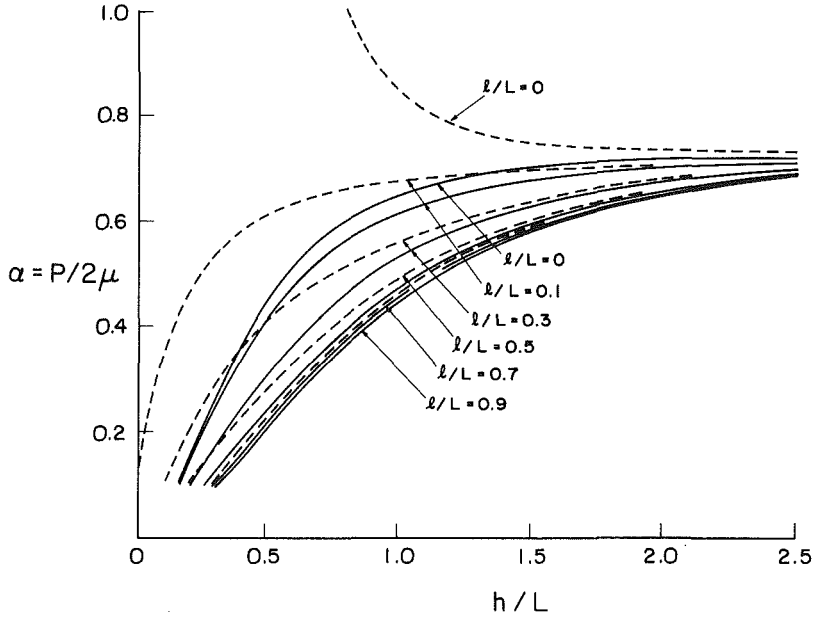


Fig. 4 Buckling load of the strip with an array of collinear cracks along the center line (ν = 0.3); solid line = antisystematic mode; dashed line = systematic mode

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

that all of the lengths are made dimensionless with respect to the crack spacing L. Poisson's ratio has the value of $\nu = 0.3$ for all calculations. When $l/L = 0$, the buckling load α achieves the value $\alpha = 0.724$ as obtained in [1]. As the values of l/L increase, the buckling load α decreases; this would tend to suggest that the cracks associated with splitting close to the surface of a layer will tend to become unstable once the cracks begin to grow. All curves with l/L different from zero are increasing functions of h/L with α approaching its limit value of $\alpha = 0.724$ for large values of h/L.

The next problem considered is the buckling of a layer with an array of cracks along its center line. Symmetric and antisymmetric modes of buckling are studied. The symmetric mode is characterized by crack opening displacements that are symmetric with respect to the center line of the layer (corresponding to Mode I—no shear) whereas the antisymmetric mode is primarily a sliding mode (corresponding to Mode II—no normal stress). The results for both cases are shown in Fig. 4 where it is seen that the buckling loads for the antisymmetric mode are lower than those for the symmetric mode; however, it is interesting to note that the difference becomes smaller as the crack length increases. In fact, for $l/L = 0.9$ the curves for both cases are almost identical to the extent that they can not be differentiated with the scale used in the plot.

The curves in Fig. 4 are an increasing function of the half width of the layer and approach the limiting value of $\alpha = 0.724$ for a large value of h/L. The only exception is the curve for the undamaged strip in the symmetric mode (dashed curve with $l/L = 0$). For this case the curve is a decreasing function of h/L and approaches the limiting value from the foregoing.

Acknowledgment

The research was supported in part by the National Science Foundation, Grant CME-8006265.

References

1 Dorris, J. F., and Nemat-Nasser, S., "Instability of a Layer on a Half Space," Technical Report No. 79-5-18, Earthquake Research and Engineering Laboratory, Dept. of Civil Engineering, Northwestern Univ., Evanston, Ill., 1979. Also in ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 304-312.

2 Horii, H., and Nemat-Nasser, S., "Instability of a Half-Space With Frictional Materials," *ZAMP*, Vol. 33, 1982, pp. 1-16.

3 Biot, M. A., *Mechanics of Incremental Deformations*, Wiley, New York, 1965.

4 Hill, R., "On Uniqueness and Stability in the Theory of Finite Elastic Strain," *J. Mech. Phys. Solids*, Vol. 5, 1957, pp. 229-241.

5 Nemat-Nasser, S., "Continuum Bases for Consistent Numerical Formulation of Finite Strains in Elastic and Inelastic Structures," *Finite Element Analysis of Transient Nonlinear Structural Behavior*, Belytschko, T., Osias, J. R., and Marcel, P. V., eds., ASME, AMD, Vol. 14, 1975, pp. 85-98.

6 Green, A. E., Rivlin, R. S., and Shield, R. T., "General Theory of Small Elastic Deformations Superposed on Finite Elastic Deformations," *Proc. Roy. Soc., London, A*, Vol. 211, 1952, pp. 128-154.

7 Sneddon, I. N., and Srivastav, R. P., "The Stress in the Vicinity of an Infinite Row of Collinear Cracks in an Elastic Body," *Proc. Roy. Soc. Edin. A*, Vol. 67, 1965, pp. 39-49.

8 Sneddon, I. N., and Lowengrub, M., *Crack Problems in the Classical Theory of Elasticity*, Wiley, New York, 1969.

9 Stahl, B., and Keer, L. M., "Vibration and Stability of Cracked Rectangular Plates," *Int. J. Solids Structures*, Vol. 8, 1972, pp. 69-91.

10 Stahl, B., and Keer, L. M., "Vibration and Buckling of a Rectangular Plate With an Internal Support," *Quar. J. Mech. and Appl. Math.*, Vol. 25, 1972, pp. 467-478.

11 Keer, L. M., and Stahl, B., "Eigenvalue Problems of Rectangular Plates With Mixed Edge Conditions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 513-520.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Approximate Analytical Model of Anchor Pull-Out Test

G. R. Miller

L. M. Keer
Mem. ASME

Department of Civil Engineering,
Northwestern University,
Evanston, Ill. 60201

The two-dimensional problem of a rigid, unbonded plate embedded in an infinite medium with cracks emanating from the edges of the plate is solved with loading conditions approximating those found in pull-out testing and other engineering applications. The analysis employs the Green's function approach of Lo, which leads to a singular integral equation that is solved numerically. Stress intensity factors are presented for several combinations of load geometry, crack length, and crack extension angle. These results are used to predict qualitively the crack propagation behavior under the assumptions of linear-elastic fracture mechanics.

Introduction

Embedded anchors have widespread applications in concrete and rock-related engineering designs. Also, an anchor pull-out test is presently being evaluated as a means of testing and characterizing the strength of concrete (see, for example, references [1–3]). Earlier analytical studies of embedded anchors have focused mainly on stress distributions and load-deflection behavior, as in the work of Selvadurai [4], Luk and Keer [5], Zhukovskii [6], and Bosakov [7]. The present study focuses on failure behavior as manifested in the propagation of cracks emanating from the edge of the anchor.

The analysis is planar and two-dimensional; the anchor is assumed to be embedded sufficiently deep to remove most boundary effects. One boundary effect, however, is approximated by including concentrated forces to simulate the loading that is often applied in testing such configurations. The Green's function approach of Lo [8] is used, which leads to a singular integral equation for the dislocation density corresponding to an extending crack with a given length and extension angle. The integral equation is solved numerically, enabling the calculation of stress intensity factors. By studying cracks of various lengths and different extension angles, critical combinations of these parameters can be estimated using linear-elastic fracture analysis, and thus the most likely failure modes predicted.

Formulation

The problem is formulated in terms of the complex potentials of Muskhelishvili [9]. The stresses and displacements can be expressed in terms of the analytic functions Φ and Ψ as

$$\sigma_{xx}+\sigma_{yy}=4\,\mathrm{Re}[\Phi(z)] \tag{1}$$

$$\sigma_{yy}-\sigma_{xx}-2i\sigma_{xy}=2(z\overline{\Phi'(z)}+\overline{\Psi(z)}) \tag{2}$$

$$2\mu(u'+iv')=\kappa\Phi(z)-z\overline{\Phi'(z)}-\overline{\Psi(z)} \tag{3}$$

where $z = x + iy$, μ is the shear modulus, and $\kappa = 3-4\nu$ for plane strain and $\kappa = (3-\nu)/(1+\nu)$ for plane stress, ν being Poisson's ratio. Primes denote differentiation with respect to z, and bars imply complex conjugation.

The anchor is modeled by a rigid unbonded plate in an infinite elastic medium loaded vertically, and failure or pullout is assumed to arise from cracking emanating from the corners of the plate (see Fig. 1). The boundary conditions corresponding to this problem are

$$2\mu\left(\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}\right)^{-}=0 \tag{4}$$

and

$$(\sigma_{yy}+i\sigma_{xy})^{+}=0 \tag{5}$$

along the plate (superscripts + and − refer to the upper and lower surfaces of the plate, respectively), while along the extending cracks the required condition is that

$$\sigma_{\phi\phi}-i\sigma_{\rho\phi}=0 \tag{6}$$

where $\sigma_{\phi\phi}$ and $\sigma_{\rho\phi}$ represent the normal and shear stresses along the crack faces. Also, the presence of the concentrated forces requires that the net traction around any sufficiently large contour containing both the plate and the concentrated loads be zero.

Before proceeding with the solution it is noted finally that equations (1) and (2) can be added to give

$$\sigma_{yy}-i\sigma_{xy}=\Phi(z)+\overline{\Phi(z)}+z\overline{\Phi'(z)}+\overline{\Psi(z)}. \tag{7}$$

Method of Solution

First, the problem of two symmetric dislocations and two symmetric concentrated forces interacting with a rigid, unbonded plate loaded vertically in an infinite body is solved. The solution is sought in the form:

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April, 1982; final revision, May, 1982. Paper No. 82-WA/APM-28.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\Phi = \Phi_D + \Phi_C + \Phi_H + \Phi_M$$
$$\Psi = \Psi_D + \Psi_C + \Psi_H + \Psi_M \tag{8}$$

where Φ_D and Ψ_D are the complex potentials due to the dislocations, Φ_C and Ψ_C are the complex potentials arising from the concentrated forces, Φ_H and Ψ_H represent the homogeneous solution for a rigidly displaced unbonded plate, and Φ_M and Ψ_M are the complex potentials required to remove the effects of the dislocations and the concentrated forces along the plate.

The singular expressions for Φ_D, Ψ_D, Φ_C, and Ψ_C are well known (see, e.g., [9, 10]) and can be written

$$\Phi_D(z) = \frac{\alpha}{(z-z_0)} + \frac{\bar{\alpha}}{(z+\bar{z}_0)},$$
$$\Psi_D(z) = \frac{\bar{\alpha}}{(z-z_0)} + \frac{\alpha}{(z+\bar{z}_0)} + \frac{\alpha\bar{z}_0}{(z-z_0)^2} - \frac{\bar{\alpha}z_0}{(z+\bar{z}_0)^2}; \tag{9}$$

$$\Phi_C = \frac{-iP}{4\pi(\kappa+1)}\left[\frac{1}{(z-z_1)} + \frac{1}{(z+\bar{z}_1)}\right],$$
$$\Psi_C = \frac{-iP\kappa}{4\pi(\kappa+1)}\left[\frac{1}{(z-z_1)} + \frac{1}{(z+\bar{z}_1)}\right] - \frac{iP}{4\pi(\kappa+1)}\left[\frac{\bar{z}_1}{(z-z_1)^2} - \frac{z_1}{(z+\bar{z}_1)^2}\right] \tag{10}$$

where in (9), $\alpha \equiv \mu e^{i\theta}([u_r] + i[u_\theta])/\pi i(\kappa+1)$; $[u_r]$ and $[u_\theta]$ represent the jump in displacements across the dislocation line. The expressions for Φ_H and Ψ_H could be written similarly, but they will arise naturally from the analysis that follows.

The contributions to the stress on the upper side of the plate resulting from the dislocations and concentrated forces can be determined directly from (7) using expressions (9) and (10) for the potentials. These contributions are

$$(\sigma_{yy} - i\sigma_{xy})^+ = \alpha\left[\frac{1}{(x-z_0)} + \frac{1}{x-\bar{z}_0} + \frac{(z_0-\bar{z}_0)}{(x+z_0)^2}\right] + \bar{\alpha}\left[\frac{1}{(x+\bar{z}_0)} + \frac{1}{(x+z_0)} + \frac{(z_0-\bar{z}_0)}{(x-\bar{z}_0)^2}\right] + \frac{iP}{4\pi(\kappa+1)}\left\{(\kappa+1)\left[\frac{1}{(x-\bar{z}_1)} + \frac{1}{(x+z_1)}\right] - \left[\frac{1}{(x-z_1)} + \frac{1}{(x+\bar{z}_1)}\right] - \left[\frac{\bar{z}_1+x}{(x+z_1)^2} - \frac{z_1-x}{(x-\bar{z}_1)^2}\right]\right\} \equiv h_\sigma(x; z_0, z_1, \alpha). \tag{11}$$

Similarly, the displacement effects on the lower side of the plate can be written:

$$2\mu\left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)^- = \alpha\left[\frac{\kappa}{(x-z_0)} - \frac{1}{(x-\bar{z}_0)} - \frac{(z_0-\bar{z}_0)}{(x+z_0)^2}\right] + \bar{\alpha}\left[\frac{\kappa}{(x+\bar{z}_0)} - \frac{1}{(x+z_0)} - \frac{(z_0-\bar{z}_0)}{(x-\bar{z}_0)^2}\right] - \frac{iP}{4\pi(\kappa+1)}\left\{(\kappa-1)\left[\frac{1}{(x-\bar{z}_1)} + \frac{1}{(x+z_1)}\right] + \kappa\left[\frac{1}{(x-z_1)} + \frac{1}{(x+\bar{z}_1)}\right] - \left[\frac{\bar{z}_1+x}{(x+z_1)^2} - \frac{z_1-x}{(x-\bar{z}_1)^2}\right]\right\} \equiv h_d(x; z_0, z_1, \alpha). \tag{12}$$

The potentials Φ_M and Ψ_M are now sought to remove the effects of (11) and (12) along the upper and lower faces of the plate, respectively. Following Muskhelishvilli (see [9], Section 120) an additional potential, Ω_M, is introduced:

$$\Omega_M(z) = \tilde{\Phi}_M(z) + z\bar{\Phi}'_M(z) + \bar{\Psi}_M(z) \tag{13}$$

where $\tilde{\Phi}(z) = \overline{\Phi(\bar{z})}$. The stress and displacement equations in terms of the new potentials become:

$$\sigma_{yy} - i\sigma_{xy} = \Phi_M(z) + \Omega_M(\bar{z}) + (z-\bar{z})\overline{\Phi'_M(z)} \tag{14}$$
$$2\mu(u' + iv') = \kappa\Phi_M(z) - \Omega_M(\bar{z}) - (z-\bar{z})\overline{\Phi'_M(z)}. \tag{15}$$

Enforcing the necessary boundary conditions to render the upper face of the plate stress free while keeping the lower face rigidly flat leads to the following Hilbert problem:

$$(\sigma_{yy} - i\sigma_{xy})^+ = \Phi_M^+ + \Omega_M^- = -h_\sigma(x; z_0, z_1, \alpha),$$
$$2\mu\left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)^- = \kappa\Phi_M^- - \Omega^+ = -h_d(x; z_0, z_1, \alpha). \tag{16}$$

The second of equation (16) is multiplied by $i/\sqrt{\kappa}$ and added to the first. This gives a new equation involving $\Phi_M - i\Omega_M/\sqrt{\kappa}$. Repeating the process with $i/\sqrt{\kappa}$ leads to a second new expression involving $\Phi_M + i\Omega_M/\sqrt{\kappa}$, and thus equation (16) is reduced to the following two equations in standard form:

$$D^+ - i\sqrt{\kappa}D^- = p_1(x)$$
$$E^+ + i\sqrt{\kappa}E^- = p_2(x) \tag{17}$$

where

$$D(z) \equiv \Phi_M(z) + \frac{i}{\sqrt{\kappa}}\Omega_M(z)$$
$$E(z) \equiv \Phi_M(z) - \frac{i}{\sqrt{\kappa}}\Omega_M(z) \tag{18}$$

and

$$p_1(x) \equiv -h_\sigma(x; z_0, z_1, \alpha) + \frac{i}{\sqrt{\kappa}}h_d(x; z_0, z_1, \alpha),$$
$$p_2(x) \equiv -h_\sigma(x; z_0, z_1, \alpha) - \frac{i}{\sqrt{\kappa}}h_d(x; z_0, z_1, \alpha). \tag{19}$$

The solutions to (17) can be written directly in terms of Cauchy integrals as

$$D(z) = \frac{X_1(z)}{2\pi i}\int_{-c}^{c}\frac{p_1(t)\,dt}{X_1^+(t)(t-z)} + A_1X_1(z),$$
$$E(z) = \frac{X_2(z)}{2\pi i}\int_{-c}^{c}\frac{p_2(t)\,dt}{X_2^+(t)(t-z)} + A_2X_2(z). \tag{20}$$

In these expressions A_1 and A_2 are constants to be determined, and $X_1(z)$ and $X_2(z)$ represent the homogeneous solutions of equation (17) given by

$$X_j(z) = (z+c)^{-\gamma_j}(z-c)^{\gamma_j-1} \tag{21}$$
$$\gamma_1 = \frac{1}{4} - i\frac{\log\kappa}{4\pi}, \quad \gamma_2 = \frac{3}{4} - i\frac{\log\kappa}{4\pi}. \tag{22}$$

It can thus be seen that upon enforcing net traction conditions and requiring single-valued displacements at infinity, the homogeneous potentials Φ_H and Ω_H become

$$\Phi_H(z) = \frac{1}{2}\left[\frac{iP(1+i\sqrt{\kappa})}{4\pi(\kappa+1)}X_1(z) + \frac{iP(1-i\sqrt{\kappa})}{4\pi(\kappa+1)}X_2(z)\right],$$
$$\Omega_H(z) = -i\frac{\sqrt{\kappa}}{2}\left[\frac{iP(1+i\sqrt{\kappa})}{4\pi(\kappa+1)}X_1(z) - \frac{iP(1-i\sqrt{\kappa})}{4\pi(\kappa+1)}X_2(z)\right]. \tag{23}$$

The integral expressions for $D(z)$ and $E(z)$ given by (20) can be evaluated through contour integration. Choosing A_1 and A_2 to ensure proper behavior at infinity the following results are obtained:

$$D(z) = \frac{1}{(1-i\sqrt{\kappa})}\{-\alpha[F_1(z;z_0) + F_1(z;\bar{z}_0) + (z_0-\bar{z}_0)G_1(z;-z_0)] - \bar{\alpha}[F_1(z;-\bar{z}_0) + F_1(z;-z_0) + (z_0-\bar{z}_0)G_1(z;\bar{z}_0)]\} + \frac{P(\sqrt{\kappa}-i)}{4\pi(\kappa+1)^2}\{(\kappa+i\sqrt{\kappa})[F_1(z;\bar{z}_1) + F_1(z;-z_1)] \tag{24}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-(1-i\sqrt{\kappa})[F_1(z;z_1)+F_1(z;-\bar{z}_1)]$$
$$-\left(1+\frac{i}{\sqrt{\kappa}}\right)(z-\bar{z}_1)[G_1(z;\bar{z}_1)+G_1(z;-z_1)]\Big\}$$

$$E(z)=\frac{1}{(1+i\sqrt{\kappa})}\Big\{-\alpha[F_2(z;z_0)+F_2(z;\bar{z}_0)+(z_0-\bar{z}_0)G_2(z;-z_0)]$$
$$-\bar{\alpha}[F_2(z;\bar{z}_0)+F_2(z;-z_0)+(z-\bar{z}_0)G_2(z;\bar{z}_0)]\Big\}$$
$$-\frac{P(\sqrt{\kappa}+i)}{4\pi(\kappa+1)^2}\Big\{(\kappa-i\sqrt{\kappa})[F_2(z;\bar{z}_1)+F_2(z;-z_1)] \tag{25}$$
$$-(1+i\sqrt{\kappa})[F_2(z;z_1)+F_2(z;\bar{z}_1)]$$
$$-\left(1-\frac{i}{\sqrt{\kappa}}\right)(z_1-\bar{z}_1)[G_2(z;\bar{z}_1)+G_2(z;z_1)]\Big\}$$

where the functions F_j and G_j are analogous to those obtained by Lo [8]:

$$F_j(z;\xi)\equiv\left[1-\frac{X_j(z)}{X_j(\xi)}\right]\frac{1}{(z-\xi)}$$
$$G_j(z;\xi)\equiv\frac{\partial}{\partial\xi}[F_j]=\left[F_j(z;\xi)+\frac{X_j(z)X_j'(\xi)}{X_j^2(\xi)}\right]\frac{1}{(z-\xi)}. \tag{26}$$

The potentials Φ_M and Ω_M can be determined using equations (18) to be

$$\Phi_M(z)=\frac{1}{2}[D(z)+E(z)]$$
$$\Omega_M(z)=-\frac{i\sqrt{\kappa}}{2}[D(z)-E(z)]. \tag{27}$$

The next step in the analysis is to replace the discrete dislocations by distributed dislocation densities. This is accomplished by replacing the complex constant α with the infinitesimal Burger's vector $b(\xi)d\xi$, where

$$b(\xi)\equiv\frac{\mu e^{i\theta}}{\pi i(\kappa+1)}\frac{\partial}{\partial\xi}\{[u_r]+i[u_\theta]\}. \tag{28}$$

The potentials Φ_D, Ψ_D, Φ_M, and Ω_M thus become integral expressions. Noting that only the first curly-bracketed term in the expressions for $D(z)$ and $E(z)$ contains α (see equations (24) and (25)), the following convenient decomposition can be made:

$$\Phi_M(z)=\Phi_{MC}(z;z_1)+\Phi_{MD}(z;\alpha,z_0). \tag{29}$$

Here the subscripts indicate the source of the effects removed by Φ_M; "*MC*" denotes concentrated force effects, and "*MD*" denotes dislocation effects. This enables the solution to the distributed dislocation problem to be written as follows:

$$\Phi(z)=\Phi_H(z)+\Phi_C(z;z_1)+\Phi_{MC}(z;z_1)+\int_a^b[\Phi_D(z;b(\xi),z_0^*)$$
$$+\Phi_{MD}(z;b(\xi),z_0^*)]d\xi \tag{30}$$

$$\Omega(z)=\Omega_H(z)+\Omega_C(z;z_1)+\Omega_{MC}(z;z_1)+\int_a^b[\Omega_D(z;b(\xi),z_0^*)$$
$$+\Omega_{MD}(z;b(\xi),z_0^*)]d\xi \tag{31}$$

with $z_0^*=\xi e^{i\theta}+\beta$,$\beta$ being a complex number specifying the location of an arbitrary point along the dislocation line, and θ being the angle of inclination of the dislocation line.

From these potentials the stresses can be obtained directly from equation (14). To model a branching crack, however, it is necessary to have an expression for the normal and shear components, $\sigma_{\phi\phi}$ and $\sigma_{\rho\phi}$, along an arbitrarily oriented line. This expression can be determined to be

$$\sigma_{\phi\phi}-i\sigma_{\rho\phi}=\Phi(z)+\overline{\Phi(z)}+e^{-2i\theta}(\Omega(\bar{z})$$
$$+(z-\bar{z})\overline{\Phi'(z)}-\overline{\Phi(z)}). \tag{32}$$

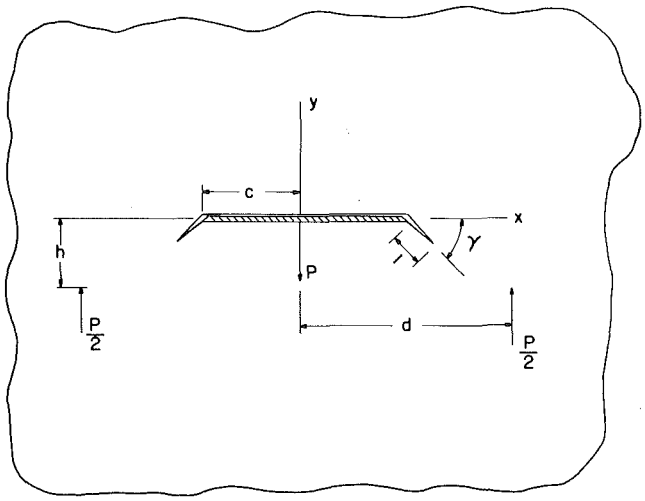


Fig. 1 Problem configuration

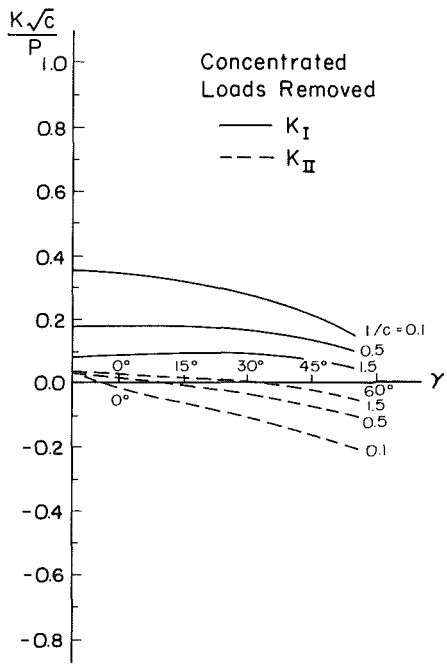


Fig. 2 Stress intensity factors as functions of extension angle: no concentrated forces

Setting equation (32) equal to zero as in equation (16) and using equations (30) and (31) the following singular integral equation is obtained:

$$\int_0^l\frac{2b(\xi)e^{-i\theta}}{t-\xi}d\xi+\int_0^l M(\hat{z}_0,\hat{z};b(\xi))d\xi+(\sigma_{\phi\phi}-i\sigma_{\rho\phi})_H$$
$$+(\sigma_{\phi\phi}-i\sigma_{\rho\phi})_C+(\sigma_{\phi\phi}-i\sigma_{\rho\phi})_{MC}=0 \tag{33}$$

where the last three terms are obtained by substituting the appropriate potentials into (32). Equation (33) is valid for the particular case of a crack of length l branching from the tip of the anchor as in Fig. 1. Thus, β has been taken to be the real number c, and $z^*=c+te^{i\theta}$. The kernel of the integral equation is composed of terms involving Φ_{MD} and Ω_{MD} and the regular part of Φ_D and Ψ_D:

$M(\hat{z}_0,\hat{z};b(\xi))$

$$=2\,\mathrm{Re}\left[\frac{b(\xi)}{(\bar{z}^*+z_0^*)}\right]+e^{-2i\theta}\left[\frac{\overline{b(\xi)}(\bar{z}^*+z_0^*)-b(\xi)(z^*-\bar{z}_0^*)}{(\bar{z}^*+z_0^*)^2}\right]$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+\Phi_{MD}(z^*;b(\xi),z_0^*)+\overline{\Phi_{MD}(z^*;b(\xi),z_0^*)}(1-e^{-2i\theta}) \quad (34)$$
$$+e^{-2i\theta}[\Omega_{MD}(\bar{z}^*;b(\xi),z_0^*)+(2t\sin\theta)\overline{\Phi'_{MD}(z^*;b(\xi),z_0^*)}].$$

Numerical Results and Conclusions

Equation (33) was solved numerically using the piecewise quadratic method of Gerasoulis [11]. It can be seen from equations (21) and (22) that the homogeneous solution (equation (23)) and terms in the kernel have an oscillating singularity at $z = \pm c$. However, in introducing the cracks, this singularity is changed and assumed to be nonoscillatory of order less than 1/2 and provides the necessary additional condition to render the solution of equation (33) unique.

The stress intensity factors, defined by

$$K_I - iK_{II} = \lim_{\xi \to l^+} \sqrt{2\pi(\xi - l)}\,(\sigma_{\phi\phi} - i\sigma_{\rho\phi}) \quad (35)$$

can be related directly to the dislocation density, $b(\xi)$, by taking the asymptotic form of equation (33). In terms of dimensionless quantities arising from the numerical scheme, the result is

$$(K_I - iK_{II})\frac{\sqrt{c}}{P} = (2\pi)^{3/2}\sqrt{\frac{c}{l}}\,e^{-i\theta}\hat{b}(1) \quad (36)$$

where

$$b(\xi) = \frac{\hat{b}(s)}{\sqrt{(1-s)s}}\frac{P}{l}; \quad s = \xi/l. \quad (37)$$

Stress intensity factors were calculated for several combinations of the parameters involving crack length (l/c), angle of extension (γ), and location of the concentrated forces (h/c and d/c).

The results of these calculations are presented in Figs. 2–5. Figures 2–4 show the values of the stress intensity factors as functions of the crack extension angle for several crack lengths. Figure 5 is a plot of the maximum value of the opening-mode stress intensity factor, K_I, versus crack length for several configurations of the concentrated forces. These load configurations correspond to those used in Figs. 2–4 and are chosen to represent: (1) relatively wide and shallow spacing of the loads ($d/c = 3.5$, $h/c = 1.0$); (2) relatively narrow and deep placement ($d/c = 2.5$, $h/c = 4.0$); and (3) the case in which the concentrated forces are not present. These three cases were found to be representative.

Observation of Figs. 2–4 shows that the maximum values of the opening-mode stress intensity factor tend to occur at points where the shear mode factor, K_{II}, is close to zero and indicates that the crack growth is governed by the openng mode (see e.g., Cotterell and Rice [12]). This criterion has recently been used by Horii and Nemat-Nasser [13] in a study of crack kinking in brittle solids where results in very close agreement with experiment were obtained. Thus, crack extension is assumed here to depend on K_I alone.

Comparison of Fig. 2 with Figs. 3 and 4 illustrates clearly the effects of introducing the concentrated forces. In Fig. 2 the values of K_I are seen to be nearly constant over a fairly large range of extension angles. The variation is seen to be significantly different, however, in Figs. 3 and 4 and much more clearly defined maximums are evident. In terms of crack propagation, this would imply that the concentrated forces tend to give the crack a more defined path of growth. Initially, small cracks would grow horizontally, and as lengthening occurs the path of propagation would be

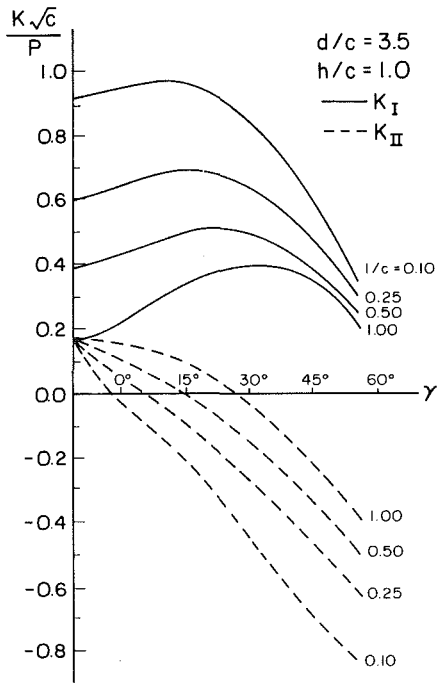


Fig. 3 Stress intensity factors as functions of extension angle: $d/c = 3.5$, $h/c = 1.0$

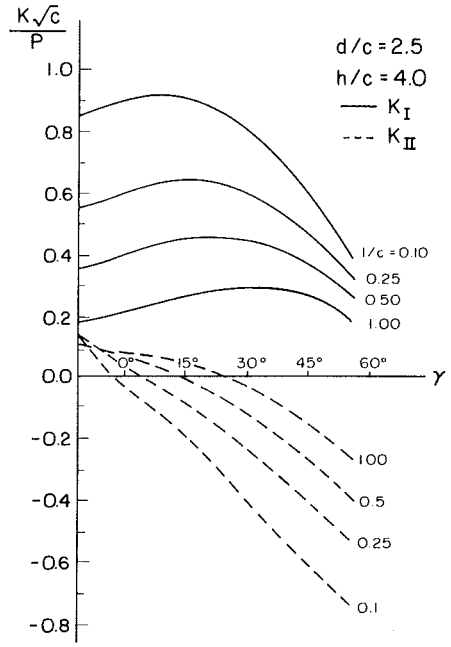


Fig. 4 Stress intensity factors as functions of extension angle: $d/c = 2.5$, $h/c = 4.0$

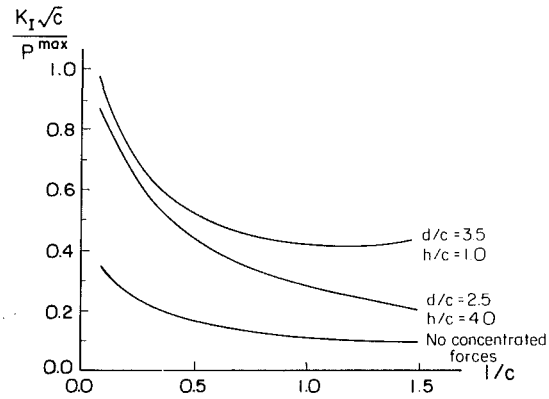


Fig. 5 Maximum opening-mode stress intensity factor as a function of crack length

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

deflected downward toward the concentrated forces. Apparently the degree of deflection depends mainly on the length of the crack, as the similarity between Figs. 3 and 4 shows. For quite different arrangements of the concentrated forces the behavior at the crack tip is altered only slightly. This was also found to be true for many other cases studied, but not shown here.

In Fig. 5, the magnitude of the maximum K_I value can be seen to increase by a factor between roughly 2.0 and 3.0 upon introducing the concentrated loads. Diminishing values of K_I with increasing crack length indicate stable crack growth, and this is initially the case for all three configurations. As the cracks lengthen, however, the placement of the loads becomes important. Wide, shallow spacing of the concentrated loads is seen to induce unstable growth in cracks that would be stable for the other configurations presented.

The results presented here should lend additional insight into the failure behavior of embedded anchors. Qualitative comparison with experiment shows fairly good agreement, although complexities not accounted for in this paper, including material nonlinearities, free surface effects, and three-dimensional effects, become important in most testing done at present, and any quantitiative model must include these factors.

Acknowledgment

The authors are grateful for support of this research from the National Science Foundation, Grant CME-8006265. The authors are also grateful for many helpful suggestions by Professor S. P. Shah.

References

1 Kierkegaard-Hansen, P., "Lok Strength," *Nordisk Betong*, No. 3, 1975, pp. 19-28.

2 Braestrup, M. W., Nielsen, M. P., Jensen, B. Chr., and Bach, F., "Axisymmetric Punching of Plain and Reinforced Concrete," Structural Research Laboratory, Technical University of Denmark, Report No. R75, 1976.

3 Jensen, B. C., and Braestrup, H. W., "Lok-Tests Determine the Compressive Strength of Concrete," *Nordisk Betong*, No. 2, 1976, pp. 9-11.

4 Selvadurai, A. P. S., "The Load-Deflexion Characteristics of a Deep Rigid Anchor in an Elastic Medium," *Géotechnique*, Vol. 26, 1976, pp. 603-612.

5 Luk, V. K., and Keer, L. M., "Stress-Analysis of a Deep Rigid Axially Loaded Cylindrical Anchor in an Elastic Medium," *International Journal for Numerical and Analytical Methods in Geomechanics*, Vol. 4, 1980, pp. 215-232.

6 Zhukovskii, I. N., "A Contact Problem for a Plate Situated in the Crack of an Infinite Body," *Prikladnaya Mekhanika*, Vol. 11, No. 11, 1975,pp. 124-128.

7 Bosakov, S. V., "Calculation of Sunken Anchor Plates of Finite Rigidity," *Prikladnaya Mekhanika*, Vol. 16, No. 3, 1980, pp. 81-87.

8 Lo, K. K., "Analysis of Branched Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1978, pp. 797-802.

9 Muskhelishvili, N. I., *Some Basic Problems in the Theory of Elasticity*, 4th ed., Noordhoff, Leyden, The Netherlands, 1954.

10 Rice, J. R., "Mathematical Analysis in the Mechanics of Fracture," *Fracture*, Vol. 2, Academic Press, New York, 1968.

11 Gerasoulis, A., "The Use of Piecewise Quadratic Polynomials for the Solution of Singular Integral Equations of the Cauchy Type," *Computers and Mathematics With Applications*, to appear.

12 Cotterell, B., and Rice, J. R., "Slightly Curved or Kinked Cracks," *Int. J. of Fracture*, Vol. 16, 1980, pp. 155-169.

13 Horii, H., and Nemat-Nasser, S., "Compression-Induced Nonplanar Crack Extension With Application to Splitting, Exfoliation, and Rockburst," *J. Geophysical Research*, to appear.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Effect of Shear and Rotary Inertia on Dynamic Fracture of a Beam or Plate in Pure Bending

C. Levy

G. Herrmann
Fellow ASME

Division of Applied Mechanics,
Department of Mechanical Engineering,
Stanford University,
Stanford, Calif. 94305

The dynamic fracture response of a long beam of brittle material subjected to pure bending is studied. If the magnitude of the applied bending moment is increased quasi-statically to a critical value, a crack will propagate from the tensile side of the beam. As an extension of previous work, the effect of shear and of rotary inertia on the moment and induced axial load at the fracturing section is included in the present analysis. Thus an improved formulation is presented by means of which the crack length, crack-tip velocity, bending moment, and axial force at the fracture section are determined as functions of time after crack initiation. It is found that the rotary effect diminishes the axial force effect and retards total fracture time whereas the shear has an opposite effect. Thus by combining the two effects (to simulate to first order the Timoshenko beam) overall fracture is retarded and better agreement with experimental data is achieved. The results also apply for plane-strain fracture of a plate in pure bending provided the value of the elastic modulus is appropriately modified.

Introduction

The present study is a continuation and extension of a previous investigation of Adeli, Herrmann, and Freund [1] and Freund and Herrmann [2] where the dynamic fracture of a long elastic, brittle beam in pure bending was studied. If the magnitude of the applied bending moment is increased to the critical value, a crack will propagate from the tensile side of the beam across a cross section. As the crack propagates across the beam, compressional waves are released from the crack tip generating an axial force at that cross section. For sufficiently large crack-tip speeds these waves were found to produce significant effects on the crack-tip motion [1]. In an attempt to further improve the model, it was suggested that a more comprehensive beam model be introduced. To this end the problem was solved by use of the Rayleigh beam model, to see the effect of rotary inertia on the load and moment at the fracturing section, and by use of the Timoshenko beam model without rotary inertia, to see the effect of shear on the load and moment at the fracturing section. These two models were used instead of the full Timoshenko beam model in order to separate the two effects. Further, the superposition of the two results could be used to detail the first-order effects of the full Timoshenko beam model. No attempt was made to offer an indepth analysis of the stress wave pattern; however, an attempt was made to model the overall influence of shear and rotary inertia by simple one-dimensional and two-dimensional models. It should be noted that these effects were not included in the previous formulation of the problem [1, 2].

The present analysis proceeds in a similar manner to that of [1]. First, it is assumed that fracture of a particular cross section may be viewed as a discontinuity in axial displacement across the original neutral axis of the beam and a discontinuity in slope of this deformed axis. Because of the symmetry condition imposed, no shear force, and hence no discontinuity in shear is assumed to exist at the fracturing cross section. On the basis of the modified Timoshenko beam theory an initial boundary-value problem is formulated to obtain relations between the bending moment and discontinuity in overall rotation, and between the axial force and the discontinuity in axial displacement of the neutral axis of the two edges of the fracturing section. The same fracture model derived in [1] and modified in [3] is then used to define a second set of relations involving the discontinuity in rotation and axial displacement with the bending moment and load. The same fracture criterion is then imposed and an efficient numerical technique is developed to solve the governing integro-differential equation. A similar, yet more condensed formulation for the Rayleigh beam model will then be given to round out this analysis.

Dynamic Beam Equations

Following the analysis of [1] the crack is assumed to cause a reduction in bending and extensional stiffness along the crack line. Because of the crack, a discontinuity in slope and in axial displacement of the neutral axis will exist. Roughly speaking, the discontinuity in slope is caused by the relative rotation of opposite sides of the fracturing section, the rotation being

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September, 1981; final revision, March, 1982. Paper No. 82-WA/APM-9.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

about a point displaced from the neutral axis, thus giving rise to a discontinuity in axial displacement of the original neutral axis at the fracturing section.

It is assumed that the beam is quasi-statically loaded by a bending moment M until fracture initiation. With the beam held at this fracture initiating moment M_f, a crack develops and begins to grow, resulting in a discontinuity in slope, denoted by $\theta(t)$, and a discontinuity in axial displacement of the neutral axis, denoted by $\delta(t)$ (see Figs. 1 and 2). Time is measured from the instant that M_f is reached, and we will only consider one half of the beam due to the assumed symmetry about the fracturing section. This condition leads us to the assumption of no shear on the fracturing section. We will denote the bending moment on the fracturing section as $M(0, t)$ and the reduced bending moment there by $m(t) = M(0, t) - M_f$.

We define the axial coordinate of the beam by x and take the fracturing section at $x = 0$. We also define the relative transverse deflection (deflection after fracture minus deflection before fracture) by $w(x, t)$. To simplify the analysis with respect to wave reflection from the ends, the beam is assumed to be (infinitely) long. We further allow the cross sections of the beam to rotate relative to the neutral axis and define the relative rotation angle (angle after fracture minus angle before fracture) by $\phi(x, t)$. To study the effect of shear only, we then use the Timoshenko beam equations, but without rotary inertia. Thus the governing equations on w and ϕ and the associated boundary and intitial conditions are

$$k\mu A(w_{,x} - \phi)_{,x} = \rho A w_{,tt} \tag{1a}$$

$$EI\phi_{,xx} + k\mu A(w_{,x} - \phi) = 0 \tag{1b}$$

$$\phi(0,t) = \theta(t)/2 \tag{1c}$$

$$k\mu A(w_{,x} - \phi) = 0 \quad \text{(no shear)} \quad \text{at} \quad x = 0 \tag{1d}$$

$$\phi(x,t) \to 0 \text{ and } \phi_{,x}(x,t) \to 0 \quad \text{as} \quad x \to \infty \tag{1e}$$

and

$$\phi(x,0) = \phi_{,t}(x,0) = 0 \quad \text{(no initial motion)} \tag{1f}$$

By combining (1a) and (1b) and using (1c)-(1f) one equation with corresponding boundary and initial conditions on the deflection w is obtained

$$w_{,xxxx} - (1/v^2)\ddot{w}_{,xx} + \kappa^4 \ddot{w} = 0 \tag{2a}$$

$$w_{,x}(0,t) = \theta(t)/2 \tag{2b}$$

$$w_{,xxx} - (1/v^2)\ddot{w}_{,x} = 0 \quad \text{at} \quad x = 0 \quad \text{(no shear)} \tag{2c}$$

$$w(x,t) \to 0 \quad \text{and} \quad w_{,x}(x,t) \to 0 \quad \text{as} \quad x \to \infty \tag{2d}$$

$$w(x,\theta) = \dot{w}(x,0) = 0 \quad \text{(no initial motion)} \tag{2e}$$

Here the comma stands for differentiation with respect to the coordinate x; the dot stands for differentiation with respect to time; $\kappa^4 = \rho A/EI$; $v^2 = k\mu/\rho$; A is the cross-sectional area; E is Young's modulus; ρ is the beam mass density; k is the shear deflection coefficient; I is the area moment of inertia; and μ is the shear modulus. It is implicit in (1a,b) and (2a) that the material properties are constant. The equations here are written for plane stress, but may be converted to plane strain by appropriate change in E.

Using the Laplace transform on (2a)-(2e), we find the relationship between $m(t)$ and $\theta(t)$. Thus

$$\theta(t) = (-2/EI\kappa^2 v)\int_0^t \{\kappa^2 v^2 e^{-\kappa^2 v^2(t-\tau)}[I_0(\kappa^2 v^2(t-\tau)) + I_1(\kappa^2 v^2(t-\kappa))]\} \times m(\tau)\,d\tau - 2m(t)/EI\kappa^2 v \tag{3}$$

where I_0 and I_1 are the modified Bessel's functions of the first kind of zero and first order, respectively.

By next considering the extensional motions of the beam as in [1], we obtain the governing equation relating the axial displacement $\delta(t)$ to the load on the fracturing section $P(t)$, (equation (10), [1])

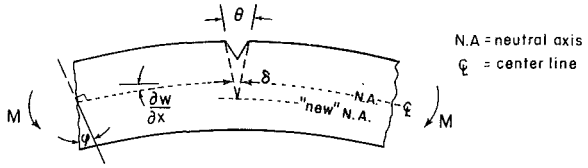


Fig. 1 Definition of terms for Timoshenko beam model



Fig. 2 Definition of reduced bending moment and load

$$\delta(t) = -(2c_0/Eh)\int_0^t P(\tau)\,d\tau \tag{4}$$

Here $c_0 = \sqrt{E/\rho}$ in plane stress and h is the height of the beam.

The Fracture Model

The analysis of this section is identical to that in [1], but with modification described by [3]. Because of the quasi-static assumption that the stress distribution on the fracturing section is related to the instantaneous crack length, net-section bending moment, and induced extensional load, this leads to the use of the stress-intensity factor of an edge-cracked strip in-plane stress or plane strain. Thus

$$K = (6M/h^{3/2})g_b(a/h) + (P/h^{1/2})g_t(a/h)^* \tag{5}$$

where g_b and g_t are dimensionless shape functions dependent on specimen geometry. If $a/h = \xi$, the ratio of the crack length to the beam height, then for $0 \le \xi \le 0.6$ g_b and g_t are given by equation (12) of [1]. For $0.6 \le \xi < 1$ g_b and g_t are given by equation (2b and 2d) of [3], i.e.,

$$g_b = \frac{0.7083 - 0.0744\xi}{(1-\xi)^{3/2}} \tag{6a}$$

$$g_t = \left(\frac{\xi}{(1-\xi)^3}\right)^{1/2}(1.7756 + 0.0528\xi) \tag{6b}$$

The Irwin relationship ($H = K^2/E$) is then used to relate the energy release rate to the stress-intensity factor. If θ and δ represent the additional rotation and translation, respectively, of one end of the strip with respect to the other due to the introduction of the crack, then we may write as in [1]

$$\theta = A_{bb}M + A_{bt}P^* \tag{7a}$$

$$\delta = A_{tt}P + A_{tb}M^* \tag{7b}$$

for our linear system. We note the compliance coefficients $A_{\lambda\mu}$ depend only on ξ. To evaluate these coefficients we next use the definition of energy release rate as being the negative of the derivative of total potential energy Φ with respect to the crack length. Thus

$$H = -\frac{\partial\Phi}{\partial a} = \frac{1}{2}\left[\frac{dA_{bb}}{da}M^2 + \frac{dA_{tb}}{da}MP + \frac{dA_{bt}}{da}PM + \frac{dA_{tt}}{da}P^2\right]^* \tag{8}$$

By comparing both forms of H, the $A_{\lambda\mu}$'s are obtained. Finally by insertion into (7a,b) the following is obtained

$$\theta = \frac{Ph^2}{D}\alpha_{bt} + \frac{6Mh}{D}\alpha_{bb} \tag{9a}$$

*See Note at end of this section.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\delta = \frac{Ph^3}{6D}\alpha_{tt} + \frac{Mh^2}{D}\alpha_{tb} \tag{9b}$$

where

$$\alpha_{tt} = \begin{cases} 1.98\xi^2 - 0.54\xi^3 + 18.65\xi^4 - 33.70\xi^5 + 99.26\xi^6 \\ -211.90\xi^7 + 436.84\xi^8 - 460.78\xi^9 + 289.98\xi^{10} \\ \\ -0.197ln(1-\xi) - 0.003\xi - (1.672 + 0.194\xi)/(1-\xi) \\ \quad + 1.672\xi/(1-\xi)^2 + 1.5633 \end{cases} \tag{10a,b}$$

$$\alpha_{tb} = \alpha_{bt} = \begin{cases} 1.98\xi^2 - 1.91\xi^3 + 16.01\xi^4 - 34.84\xi^5 + 83.93\xi^6 \\ -153.65\xi^7 + 256.72\xi^8 - 244.67\xi^9 + 133.55\xi^{10} \\ \\ \sqrt{\xi}[0.0078 - 0.1873/(1-\xi) \\ \quad + 0.5796/(1-\xi)^2] - 0.2ln\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) - 0.1021 \end{cases} \tag{11a,b}$$

$$\alpha_{bb} = \begin{cases} 1.98\xi^2 - 3.28\xi^3 + 14.43\xi^4 - 31.26\xi^5 + 63.56\xi^6 \\ -103.36\xi^7 + 147.52\xi^8 - 127.69\xi^9 + 61.504\xi^{10} \\ \\ -0.0055ln(1-\xi) + 0.944/(1-\xi) + 0.201/(1-\xi)^2 \\ \quad - 0.4178 \end{cases} \tag{12a,b}$$

Equations (10*a*, 11*a*, 12*a*) are applicable for $0 \le \xi \le 0.6$ and (10*b*, 11*b*, 12*b*) are applicable between $0.6 \le \xi < 1$. It must be noted that equations (7*a*) and (9*a*) may be written as shown since at the fracturing section the cross section is perpendicular to the neutral axis due to the zero shear condition.* By equating (3) and (4) with (9*a*, *b*), respectively, we obtain the two equations needed for the reduced bending moment $m(t)$ and the axial force $P(t)$. By defining

$$t_1 = c_0 t/h; \quad m_1(t) = m(t)/M_f; \quad P_1(t) = P(t)h/6M_f; \quad I = h^3/12;$$

$$\bar{\sigma} = \kappa^2 v^2 \frac{h}{c_0} = \sqrt{3}k/(1+\nu) \tag{13}$$

where ν is Poisson's ratio, then these two equations become in nondimensional form

$$\bar{\sigma}^{1/2}/(\sqrt{2}(3)^{5/4}) \cdot \left\{ \int_0^{t_1} e^{-\bar{\sigma}(t_1-\tau_1)} \{ I_0(\bar{\sigma}[t_1-\tau_1]) + I_1(\bar{\sigma}[t_1-\tau_1]) \} \right.$$

$$\left. \times\ m_1(\tau_1)d\tau_1 + \frac{m_1(t)_1}{\bar{\sigma}} \right\} = -[P_1\alpha_{bt} + (1+m_1)\alpha_{bb}] \tag{14}$$

$$\int_0^{t_1} P_1(\tau_1)d\tau_1 = -[P_1\alpha_{tt} + (1+m_1)\alpha_{tb}] \tag{15}$$

It should be noted that these two integral equations are implicit functions of ξ and that the actual motion of the crack is determined by imposing a crack propagation criterion identical to that used in [1].

Note

Since the model investigation includes the effect of shear, all starred (*) equations should be modified to account for this to include another equation $w = w(M, P, \text{shear})$. However, since mode *I* fracture predominates, an order anlaysis shows that the equations as written are correct to leading order. It is to be noted that when the shear is zero, as it is at the fracturing cross section, these equations become exact.

The Crack Propagation Condition

The crack propagation criterion will be briefly discussed here. In order for crack propagation to occur it is assumed that the stress-intensity factor reaches a critical value, K_f. Since we assume that fracture begins from the state of zero crack length, an infinite bending moment M_f would be required for fracture initiation. This difficulty is circumvented by introducing an effective crack length for fracture initiation a_f, which is defined by

$$K_f = 6M_f g_b(a_f/h)/h^{3/2} \tag{16}$$

where M_f is the critical bending moment previously discussed and K_f is the measured critical value of the stress-intensity factor for a precracked specimen. Thus by equating (16) and (5) we have the criterion needed for fracture when $a > a_f$. For $0 \le a < a_f$ equation (29) of [1] is used since P is negligible for very small values of a. In addition, if we take into account the effect of crack-tip speed as discussed in [3], the fracture condition is then given by equation (31) of [1] and rewritten here in its nondimensional form for completeness.

$$1 + m_1 + P_1 g_t/g_b = G(a)/[1 - (\gamma\dot{a}/c_0)^q] \tag{17}$$

with

$$G(a) = \begin{cases} 1 - 0.2032(a/a_f)^{2/3} & 0 \le a < \frac{49}{16}a_f \\ \\ g_b(a_f/h)/g_b(a/h) & a \ge \frac{49}{16}a_f \end{cases} \tag{18a,b}$$

and $\gamma \cong 4$, $q = 2$ or 3. With (14), (15), and (17) we have three highly nonlinear integro-differential equations for which an efficient algorithm was devised to obtain $a(t)$, $m_1(t)$, and $P_1(t)$.

Solution Procedure

Since all three integro-differential equations must be solved simultaneously for each time step, explicit formulas were derived for P_1 and m_1 at any time step n which are denoted by $P_1^{(n)}$ and $m_1^{(n)}$, respectively. For the initial step when $a < 49/16\, a_f$, P_1 and m_1 are denoted by $\bar{P}_1$ and $\bar{m}_1$, respectively. Using the trapezoidal rule to evaluate the integral and a suitable time step leads to

$$\bar{P}_1 = \frac{\left(\frac{1}{\lambda\bar{\sigma}^{1/2}} + \frac{\bar{t}_1\bar{\sigma}^{1/2}}{2\lambda}\right)\alpha_{tb}}{\alpha_{tb}^2 - (\bar{t}_{1/2} + \alpha_{tt})\left(\frac{1}{\lambda\bar{\sigma}^{1/2}} + \frac{\bar{t}\bar{\sigma}^{1/2}}{2\lambda} + \alpha_{bb}\right)} \tag{19a}$$

$$\bar{m}_1 = \frac{\alpha_{bb}(\bar{t}_{1/2} + \alpha_{tt}) - \alpha_{tb}^2}{\alpha_{tb}^2 - (\bar{t}_{1/2} + \alpha_{tt})\left(\frac{1}{\lambda\bar{\sigma}^{1/2}} + \frac{\bar{t}_1\bar{\sigma}^{1/2}}{2\lambda} + \alpha_{bb}\right)} \tag{19b}$$

where $\bar{t}_1$ is the value of t_1 when $a = 49/16\, a_f$ and $\lambda = \sqrt{2}(3)^{5/4}$. For $n \ge 1$

$$P_1^{(n)} = \frac{\left(\frac{1}{\lambda\bar{\sigma}_{1/2}} + \frac{\Delta t_1\bar{\sigma}^{1/2}}{2\lambda} + \alpha_{bb}\right)(\alpha_{tb} + A^{(n+1)}) - \alpha_{tb}(\alpha_{bb} + B^{(n-1)})}{\alpha_{tb}^2 - \left(\frac{\Delta t_1}{2} + \alpha_{tt}\right)\left(\frac{1}{\lambda\bar{\sigma}^{1/2}} + \frac{\Delta t_1\bar{\sigma}^{1/2}}{2\lambda} + \alpha_{bb}\right)} \tag{20a}$$

*See Note at end of this section.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$m_1^{(n)} = \frac{(\alpha_{bb} + B^{(n-1)})\left(\frac{\Delta t_1}{2} + \alpha_{tt}\right) - \alpha_{tb}(\alpha_{tb} + A^{(n-1)})}{\alpha_{tb}^2 - \left(\frac{\Delta t_1}{2} + \alpha_{tt}\right)\left(\frac{1}{\lambda \bar{\sigma}^{1/2}} + \frac{\Delta t_1 \bar{\sigma}^{1/2}}{2\lambda} + \alpha_{bb}\right)} \quad (20b)$$

where Δt_1 is the time step used and

$$A^{(n)} = A^{(n-1)} + \Delta t_1 P_1^{(n-1)} \quad n > 1 \quad (21a)$$

$$B^{(n)} = \frac{\bar{\sigma}^{1/2}}{\lambda}\left\{(\omega_1 + \bar{\omega}_1)\sum_{i=0}^{n-1} m_1^{(i)} e^{-\bar{\sigma}\Delta t_1 (n-i)} \times [I_0(\bar{\sigma}\Delta 1, (n-i)) + I_1(\bar{\sigma}\Delta t_1 (n-i))]\right\} \quad n \geq 1 \quad (21b)$$

The restrictions on (21*a*,*b*) are that $A^{(1)} = \bar{t}_1 + \Delta t_1/2\ \bar{P}$, $P_1^{(0)} = 0$, and when $i = 0$ $m_1^{(0)} = \bar{m}_1$ and $\bar{\omega}_1 = \bar{t}_1/2$. For $i \geq 1$ $\bar{\omega}_1 = \omega_1 = \Delta t_1/2$. I_0 and I_1 were computed using their asymptotic forms as found in Abramowitz and Stegun [4]. Since these are explicit relations for the load and moment, respectively, at time step n, the velocity may then be obtained from equation (17) by direct substitution. Further, to solve this first-order differential equation the predictor-corrector scheme discussed in [1] was used.

The Rayleigh Beam Model

The same problem of a single edge-notch specimen under applied remote bending was studied using the Rayleigh beam model to determine what effect rotary inertia would have on the bending moment and induced loading at the fracture section. Again we examine the effects of the crack on the neutral axis of the beam and define $m(t)$, $\theta(t)$, and $\delta(t)$ as in the previous case. We consider only one-half of the beam with the beam axial coordinate being x, and $x = 0$ being taken at the fracturing section (see Fig. 3). The relative transverse deflection is similarly defined by $W(x, t)$. Because of symmetry the assumption of zero shear is again used. Thus the governing equation on W and the associated boundary and initial conditions are

$$W_{,xxxx} + \kappa^4 \ddot{W} - \frac{1}{c^2}\ddot{W}_{,xx} = 0 \quad (22a)$$

$$W_{,x}(0, t) = \theta(t)/2 \quad (22b)$$

$$W_{,xxx} - \frac{1}{c^2}\ddot{W}_{,x} = 0 \quad \text{at} \quad x = 0 \quad \text{(no shear)} \quad (22c)$$

$$W(x, t) \to 0 \quad \text{and} \quad W_{,x}(x, t) \to 0 \quad \text{as} \quad x \to \infty \quad (22d)$$

$$W(x, 0) = 0 \quad \text{and} \quad \dot{W}(x, 0) = 0 \quad (22e)$$

Here $c^2 = EI/I_\rho$, where I_ρ is the mass moment of inertia about the neutral axis per length of beam ($= \rho I$ for our case). Henceforth only new symbols will be defined.

By use of the standard Laplace transform on (22*a*)–(22*c*) we find the relationship governing $\theta(t)$ in terms of $m(t)$. Thus

$$\theta(t) = -2c\left[\int_0^t e^{-\kappa^2 c^2 (t-\tau)}\left\{I_0(\kappa^2 c^2 [t-\tau]) + \int_0^{\kappa^2 c^2 (t-\tau)} I_0(u)du\right\} \times m(\tau)d\tau\right]/EI \quad (23)$$

This relation replaces equation (3) in the fracture model, all else remaining unchanged. Thus by using the first four relations of (13) and redefining

$$\bar{\sigma} = \kappa^2 c^2 \frac{h}{c_0} = 2\sqrt{3} \quad (24)$$

equation (23), nondimensionalized, leads to

$$\frac{1}{3}\int_0^{t_1} m_1(\tau_1)[\bar{e}^{\bar{\sigma}(t_1-\tau_1)}\left\{I_0(\bar{\sigma}[t_1-\tau_1]) + \int_0^{\bar{\sigma}(t_1-\tau_1)} I_0(u)du\right\}]d\tau = -[P_1\alpha_{bt} + (1+m_1)\alpha_{bb}] \quad (25)$$

where (15) remains unaffected. The same crack-propagation criterion is invoked allowing us to use (17) and (18).

Using the same type of numerical procedure previously described gives explicit values for m_1, P_1, and $\dot{a}/c_0$ at each time step. The inner integral in (25) was evaluated using one-tenth of the time step of the outer integral. The predictor-corrector scheme discussed in [1] was used with proper modification of $A^{(n)}$ and $B^{(n)}$ for this case.

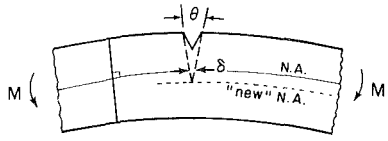


Fig. 3 Definition of terms for the Rayleigh beam model

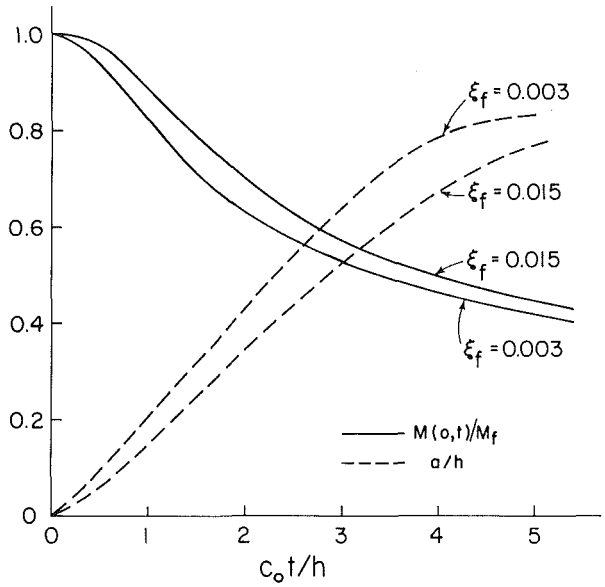


Fig. 4 Dimensionless crack length and moment versus dimensionless time for Timoshenko beam without rotary inertia effect; $\bar{\sigma} = 1.2$

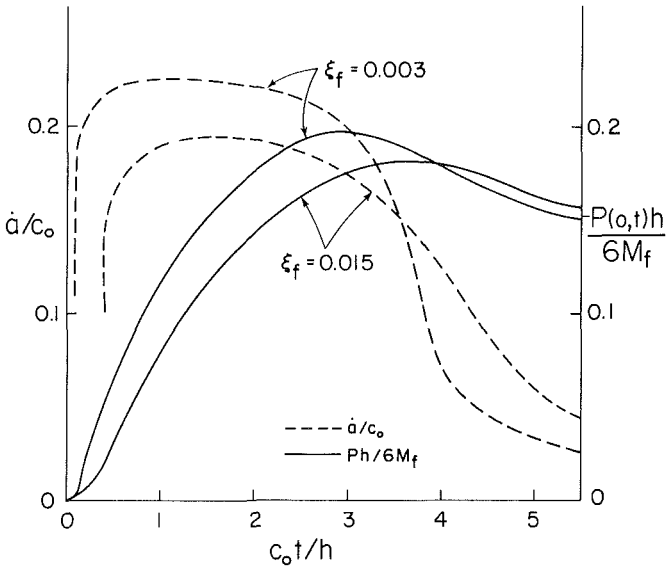


Fig. 5 Dimensionless crack speed ($\dot{a}/c_0$) and axial load on fracture face ($Ph/6M_f$) versus dimensionless time ($c_0 t/h$) for Timoshenko beam without rotary inertia effect; $\bar{\sigma} = 1.2$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

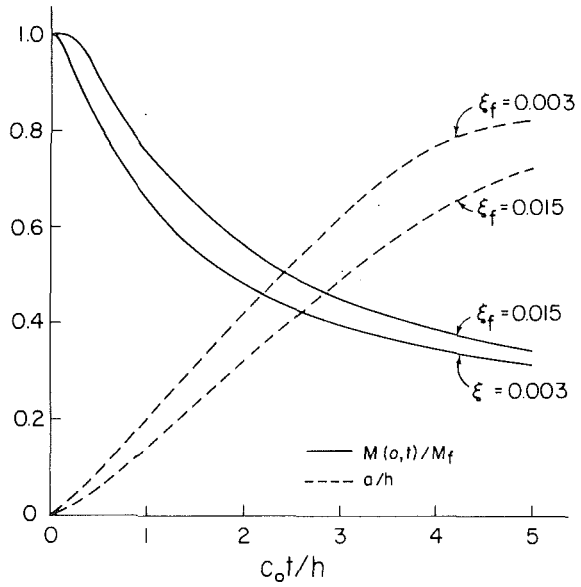


Fig. 6 Dimensionless crack length and moment versus dimensionless time for the Rayleigh model

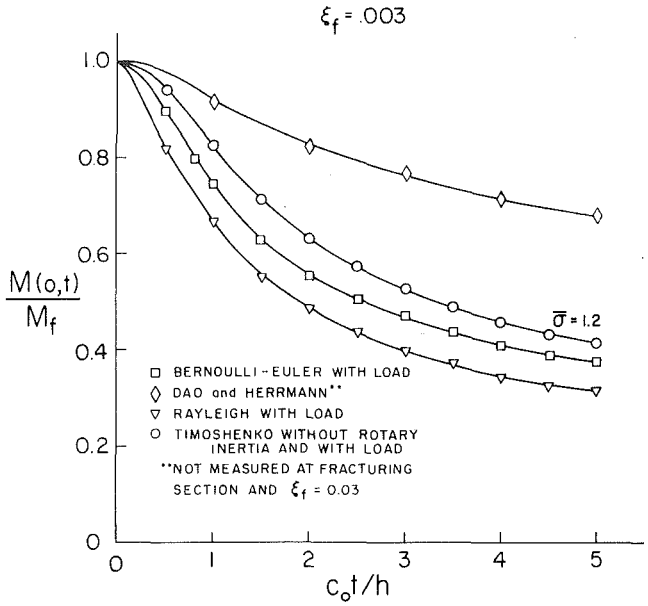


Fig. 8 Dimensionless moment versus dimensionless time for the models listed

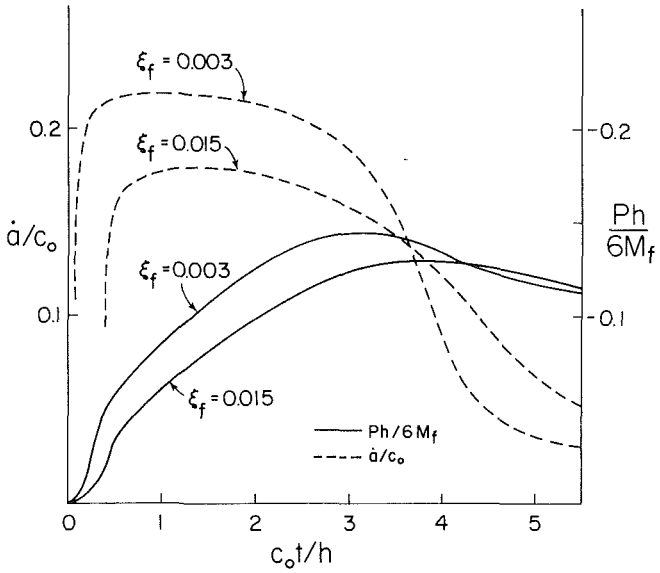


Fig. 7 Dimensionless crack speed ($\dot{a}/c_0$) and dimensionless load on fracture face ($Ph/6M_f$) versus dimensionless time ($c_0 t/h$) for the Rayleigh model

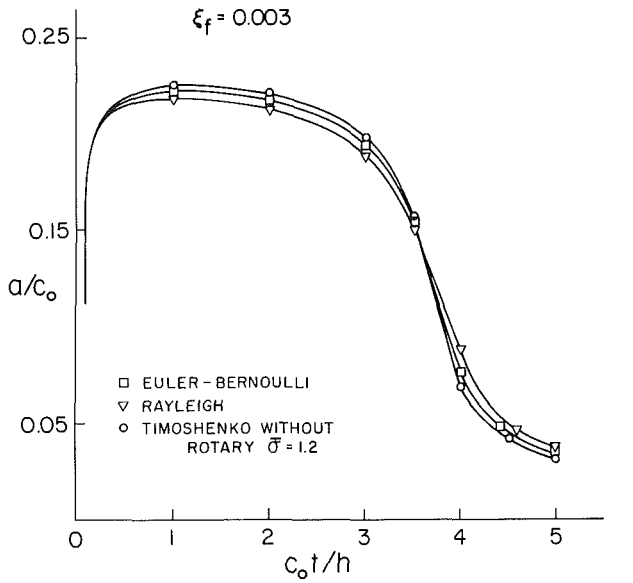


Fig. 9 Dimensionless speed versus dimensionless time for the models listed

Results and Discussion

By the algorithm in [1] with proper modification for the problems at hand, representative results were obtained for the chosen values of $\xi_f (= a_f/h)$ of 0.003 and 0.015, $q = 2$, and $\gamma = 4$. These results are plotted in terms of the dimensionless crack length a/h, bending moment $M(0, t)/M_f$, induced load $Ph/6M_f$, and velocity $\dot{a}/c_0$ versus dimensionless time $c_0 t/h$. Since $\bar{\sigma}$ is a parameter for the Timoshenko model it was taken equal to 1.2, simulating a rectangular PMMA beam.

(a) Timoshenko Model. Plotted in Fig. 4 are the dimensionless crack length and the moment on the fracturing section versus dimensionless time. In Fig. 5 the dimensionless induced load and the velocity are plotted. Several trends similar to those found for the Euler-Bernoulli model also occur here. Larger ξ_f means retardation of crack propagation, increasing the fracture section bending moment, increasing the induced load for approximately 65–80 percent of the crack length before decreasing it, and lowering the crack speed for approximately 60–75 percent of the crack length before increasing it.

In comparison with [3] we note that the shear tends to speed up crack extension, to increase the bending moment, to decrease the induced load, and to increase the velocity for over 70 percent of the height of the beam before decreasing it (see Figs. 8 and 9), if the same crack propagation criterion is kept. Further decrease in ξ_f may result in closer agreement of the two models. Similarly the experimental results plotted for a/h in Fig. 1 of [1] are in good agreement with those shown in Fig. 4.

(b) Rayleigh Model. Plotted in Fig. 6 are the dimensionless crack length and the moment on the fracturing section versus dimensionless time. Also plotted and shown in Fig. 7 are the dimensionless induced load and the velocity. We note the same trends, with respect to those of the Euler-Bernoulli model, in the crack length, bending moment, in-

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

duced load, and velocity as previously discussed. However, the effect of rotary inertia appears to retard crack growth, to decrease the bending moment on the fracturing section, to increase the induced load, and to reduce the velocity for about 70 percent of the height of the beam (see Figs. 8 and 9). We further note the good agreement of a/h in Fig. 6 with that of Fig. 1 of [1]. We also may conjecture, as with the Timoshenko model, that as ξ_f becomes smaller, the Euler-Bernoulli and Rayleigh model data coalesce.

(c) **Comparison of Models.** Plotted in Fig. 8 are the moment at the fracturing section as computed by the Euler-Bernoulli, the Rayleigh, and the Timoshenko beam without rotary inertia models as well as the data of Dao and Herrmann [5] plotted in Fig. 2 of [1]. We note that all three models exhibit the same characteristics as those of Dao and Herrmann but lower in value. The higher values of Dao and Herrmann occur because the moment measured is not at the fracturing cross section and the crack length ξ_f is larger; hence a direct comparison cannot be carried out. Also Figs. 8 and 9 clearly indicate that the results of the Euler-Bernoulli model are bracketed by our two models. Thus by "adding" the Rayleigh beam model to the Timoshenko beam without rotary inertia, one finds that the overall results are either close to the Euler-Bernoulli model or favor, very slightly, the Rayleigh beam model results. This "addition" may be considered a first-order approximation to the Timoshenko beam with rotary inertia, and if so, the Timoshenko beam with rotary inertia will give results close to those of the Euler-Bernoulli model. This means that the Euler-Bernoulli theory provides us with an adequate model for an infinitely long notched beam under external bending moments, and that wave reflections for the ends must be investigated for the finite-length beam in order to explain the final fracture of the beam. It is planned to examine this feature in the near future.

Acknowledgment

This research was supported in part by the U.S. Army Research Office under Grant DAAG29-78-G-0086 to Stanford University.

References

1 Adeli, H., Herrmann, G., and Freund, L. B., "Effect of Axial Force on Dynamic Fracture of a Beam or Plate in Pure Bending," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, 1977, pp. 647–651.

2 Freund, L. B., and Herrmann, G., "Dynamic Fracture of a Beam or Plate in Plane Bending," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, 1976, pp. 112–116.

3 Levy, C., and Herrmann, G., "On the Effect of Axial Force on Dynamic Fracture of a Beam or a Plate in Pure Bending," to be published.

4 Abramowitz, M., and Stegun, L. A., *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series 55, 9th printing, 1970, p. 378.

5 Dao, K. C., and Herrmann, G., "Crack Propagation and Stress Waves in PMMA Beams During Fracture in Bending," *Proceedings of an International Conference on Fracture Mechanics and Technology* held in Hong Kong, Mar. 1977, Vol. II, pp. 1265–1279, Sih, G. C., and Chow, C. L., eds., Sijthoff and Noordhoff, The Netherlands.

6 Sih, G. C., *Handbook of Stress-Intensity Factors*, Institute of Fracture and Solid Mechanics, Lehigh University, Bethlehem, Pa., 1973, pp. i–iii.

7 Tada, H., Paris, P., and Irwin, G., *The Stress Analysis of Cracks Handbook*, Del Research Corp., Hellertown, Pa., 1973, pp. 1–12.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Extension and Torsion of Elastic Bars With Initial Twist

R. T. Shield
Professor and Head,
Department of Theoretical
and Applied Mechanics,
University of Illinois
at Urbana-Champaign,
Urbana, Ill. 61801
Mem. ASME

In contrast to straight bars, elastic bars with initial twist exhibit coupling between tension and torsion. The coupling is examined for pretwisted bars composed of homogeneous isotropic material under the assumption that each section undergoes the same deformation. The expressions for the axial force and torsional moment are transformed by use of the governing equations and boundary conditions to forms that are particularly useful for rods with small amounts of pretwist. The elliptical section with center on the axis of pretwist and the off-centered circular section are treated in order to illustrate the coupling between extension and torsion caused by the initial twist.

Introduction

In this paper we consider the extension and torsion of an elastic bar with initial twist. The section of the bar perpendicular to its axis is the same along the bar but rotates uniformly with axial distance. The bar is composed of homogeneous isotropic material and under the assumption that each section undergoes the same deformation, the problem reduces to the determination of three displacement functions of two variables of position in the section. The functions satisfy three equations of equilibrium in the cross section and three boundary conditions to ensure that the (curved) lateral surface of the twisted bar is free from surface traction.

The present work extends the work of Okubo [1, 2], who derived the governing equations for extension and torsion of a twisted bar in forms particularly useful for bars with small amounts of pretwist. Okubo [3] has also provided a formulation that is more convenient for bars with large amounts of pretwist, such as a closely coiled helical spring.

After formulation of the problem, the reciprocal theorem is used to motivate transformations of the expressions for the axial force and twisting moment through the use of the governing equations and boundary conditions. The alternate expressions are particularly useful when approximate solutions for small amounts of pretwist are being found. In the same way, alternate expressions for the bending moments on a section are derived; the expressions are used to determine constants appearing in the solutions from the requirement that the bending moments on a section vanish.

The extension and torsion of a rod with small initial twist are treated separately, because superposition can be used for combined extension and torsion. When the axis of pretwist passes through the centroid of the cross section and the section has two axes of symmetry, the axial load L can be found to $O(\alpha^2)$, where α is the pretwist per unit length, from the geometry of the section and knowledge of the classical torsional rigidity of the section. The moment M is known to $O(\alpha)$ and can be found to $O(\alpha^2)$ when the solution to an auxiliary plane strain problem is known.

When the axis of pretwist is the line of centroids, the bar tends to unwind when it is extended and to twist further when it is compressed. Similarly, on twisting, the bar tends to contract or extend accordingly as the torsion increases or decreases the amount of pretwist. When the centroids of the cross sections do not lie on the axis of pretwist, whether the rod will tend to unwind or not on extension depends on the shape of the cross section and the value of Poisson's ratio.

The elliptical section with center on the axis of pretwist and the off-centered circular section are treated in order to illustrate the theory. For elongated elliptical sections the initial twist can cause an appreciable increase in the torsional rigidity for torsion with no extension and for torsion with no axial load. Chu [4] has shown by an approximate analysis that the torsional rigidity of bars with elongated sections is increased by the pretwist and has confirmed the theory by experiment for torsion under no axial load. The axial stiffness is also reduced appreciably, even for small amounts of pretwist, if a twisting moment is not applied to prevent the rod from unwinding.

For the off-centered circular section, the results for small amounts of pretwist show that the bar tends to twist further when it is extended and to unwind when it is compressed. However for the off-centered elliptical section with the minor axis in the radial direction, the spiral rod will tend to unwind when extended if the eccentricity of the section is large enough.

The coupling between extension and torsion of bars with initial twist has been examined recently by Rosen [5] and Hodges [6] (see also the discussion by these authors [7]). Both use variational methods but in common with an earlier treatment by Washizu [8], they do not use general enough forms for the displacements. Aside from the effect of the pretwist on the transverse displacements, they do not allow for the Poisson contraction on extension. With their ap-

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January, 1982; final revision, May, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

proach, the condition of zero traction on the (curved) lateral surface of the bar cannot be satisfied.

Knowles and Reissner [9] investigated the extension and torsion of a pretwisted strip of narrow rectangular section by treating the strip as a thin helicoidal shell. They obtained explicit results for the twisting moment and axial force correct to the fourth power of the amount of pretwist. Their results extend the analysis of Chu [4] for rectangular strips and also show that the axial stiffness for extension with no moment is reduced appreciably by the pretwist. Reissner and Wan [10] continued the shell approach of [9] and provided results for all values of the pretwist for the narrow rectangular strip. They also treated helicoidal shells with the center of the section off the axis of the pretwist and gave perturbation analyses for very small and very large amounts of pretwist.

Torsion and Extension of a Cylinder With Initial Twist

We consider a twisted beam or a spiral rod and denote the initial twist per unit length by α. The z-axis is taken to be the axis of twist of the rod and this is not necessarily the line of centroids of the cross sections, which is a helix in general. The coordinates ξ, η are defined through

$$\xi = x\cos\alpha z + y\sin\alpha z, \quad \eta = -x\sin\alpha z + y\cos\alpha z. \tag{1}$$

The ξ and η directions then have the same orientation with regard to the cross section at each section, and the lateral surface will be given by an equation of the form

$$F(\xi,\eta) = 0. \tag{2}$$

For torsion and extension of a long rod each section will undergo the same deformation so we set

$$\begin{aligned} u_\xi &= -\theta\eta z - \nu e\xi + u(\xi,\eta), \\ u_\eta &= \theta\xi z - \nu e\eta + v(\xi,\eta), \\ u_z &= ez + w(\xi,\eta), \end{aligned} \tag{3}$$

where (u_ξ, u_η, u_z) are the displacements in the ξ,η, and axial directions. The extension of the rod is e, the additional twist per unit length is θ, and ν is Poisson's ratio. It can be shown that (3) represents the most general displacement field with strains associated with the ξ,η, and z directions which are functions of ξ and η only, apart from an arbitrary rigid displacement. The functions u,v,w are to be chosen to satisfy equilibrium and to clear the lateral surface from traction.

The displacements in the x,y-directions are

$$\begin{aligned} u_x &= -\theta yz - \nu ex + u\cos\alpha z - v\sin\alpha z, \\ u_y &= \theta xz - \nu ey + u\sin\alpha z + v\cos\alpha z. \end{aligned}$$

After calculating the strains and substituting in the stress-strain relation for an isotropic elastic material, we find that the equilibrium equations reduce to

$$(\lambda+2\mu)\frac{\partial^2 u}{\partial\xi^2} + \mu\frac{\partial^2 u}{\partial\eta^2} + (\lambda+\mu)\frac{\partial^2 v}{\partial\xi\partial\eta} + \alpha(\lambda+\mu)\frac{\partial}{\partial\xi}(Dw) - \mu\alpha^2(u + 2Dv - D^2u) = 0, \tag{4}$$

$$(\lambda+2\mu)\frac{\partial^2 v}{\partial\eta^2} + \mu\frac{\partial^2 v}{\partial\xi^2} + (\lambda+\mu)\frac{\partial^2 u}{\partial\xi\partial\eta} + \alpha(\lambda+\mu)\frac{\partial}{\partial\eta}(Dw) - \mu\alpha^2(v - 2Du - D^2v) = 0, \tag{5}$$

$$\mu\left(\frac{\partial^2 w}{\partial\xi^2} + \frac{\partial^2 w}{\partial\eta^2}\right) + \alpha(\lambda+\mu)D\left(\frac{\partial u}{\partial\xi} + \frac{\partial y}{\partial\eta}\right) + \alpha^2(\lambda+2\mu)D^2 w = 0 \tag{6}$$

throughout the cross section A of the helicoid. The symbol D denotes the operator

$$D = \eta\frac{\partial}{\partial\xi} - \xi\frac{\partial}{\partial\eta},$$

and λ and μ are Lame's constants. The normal to the surface (2) is in the direction of the vector

$$\left(\cos\alpha z\frac{\partial F}{\partial\xi} - \sin\alpha z\frac{\partial F}{\partial\eta}, \sin\alpha z\frac{\partial F}{\partial\xi} + \cos\alpha z\frac{\partial F}{\partial\eta}, \alpha DF\right)$$

referred to the x,y,z-directions. It is then found that in order for the surface (2) to be free from traction we need

$$\left\{(\lambda+2\mu)\frac{\partial u}{\partial\xi} + \lambda\frac{\partial v}{\partial\eta}\right\}\frac{\partial F}{\partial\xi} + \mu\left(\frac{\partial u}{\partial\eta} + \frac{\partial v}{\partial\xi}\right)\frac{\partial F}{\partial\eta} + \alpha\left\{\lambda\frac{\partial F}{\partial\xi}Dw + \mu\left(\frac{\partial w}{\partial\xi} - \theta\eta\right)DF\right\} - \mu\alpha^2(v - Du)DF = 0, \tag{7}$$

$$\left\{(\lambda+2\mu)\frac{\partial v}{\partial\eta} + \lambda\frac{\partial u}{\partial\xi}\right\}\frac{\partial F}{\partial\eta} + \mu\left(\frac{\partial u}{\partial\eta} + \frac{\partial v}{\partial\xi}\right)\frac{\partial F}{\partial\xi} + \alpha\left\{\lambda\frac{\partial F}{\partial\eta}Dw + \mu\left(\frac{\partial w}{\partial\eta} + \theta\xi\right)DF\right\} + \mu\alpha^2(u + Dv)DF = 0, \tag{8}$$

$$\mu\left\{\left(\frac{\partial w}{\partial\xi} - \theta\eta\right)\frac{\partial F}{\partial\xi} + \left(\frac{\partial w}{\partial\eta} + \theta\xi\right)\frac{\partial F}{\partial\eta}\right\} + \alpha\left\{\left[Ee + \lambda\left(\frac{\partial u}{\partial\xi} + \frac{\partial v}{\partial\eta}\right)\right]DF - \mu(v - Du)\frac{\partial F}{\partial\xi} + \mu(u + Dv)\frac{\partial F}{\partial\eta}\right\} + \alpha^2(\lambda+2\mu)DwDF = 0 \tag{9}$$

on the boundary C of the cross section A, where E is Young's modulus. Equations (4)–(6) are in agreement with the equations derived in [1]. Note that when θ has the same sign as α, the torsion increases the pretwist of the bar, whereas in [1], the pretwist k corresponds to $-\alpha$ of this paper.

The strain components referred to fixed axes in the ξ, η, and z-directions are given by

$$e_{\xi\xi} = \frac{\partial u}{\partial\xi} - \nu e, \quad e_{\eta\eta} = \frac{\partial v}{\partial\eta} - \nu e, \quad 2e_{\xi\eta} = \frac{\partial u}{\partial\eta} + \frac{\partial v}{\partial\xi},$$

$$2e_{\xi z} = \frac{\partial w}{\partial\xi} - \theta\eta - \alpha(v - Du), \quad 2e_{\eta z} = \frac{\partial w}{\partial\eta} + \theta\xi + \alpha(u + Dv), \quad e_{zz} = e + \alpha Dw. \tag{10}$$

The Axial Force and Twisting Moment

The strain energy U per unit length of the rod (in the axial direction) is a quadratic in e and θ, and the axial force L and the twisting moment M on a section are the derivatives of U with respect to e and θ. It follows that L and M are of the form

$$L = le + k\theta, \quad M = ke + m\theta.$$

The stiffness matrix can also be seen to be symmetric because of reciprocity (see also [9]). The coefficients $l(\alpha)$ and $m(\alpha)$ will be even in α because the extensional stiffness under zero twist and the torsional modulus under zero extension cannot depend on the sign of α. The coupling coefficient $k(\alpha)$ must

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

change sign when the sign of α is changed and therefore must be an odd function of α. The axial force L and the twisting moment M are given by

$$L = \mathrm{E}eA + \iint_A \left\{\lambda\left(\frac{\partial u}{\partial \xi} + \frac{\partial v}{\partial \eta}\right) + \alpha(\lambda+2\mu)Dw\right\} d\xi d\eta, \quad (11)$$

$$M = \mu\iint_A \left\{\xi\left(\frac{\partial w}{\partial \eta} + \theta\xi\right) - \eta\left(\frac{\partial w}{\partial \xi} - \theta\eta\right) + \alpha[\xi(u+Dv) + \eta(v-Du)]\right\} d\xi d\eta, \quad (12)$$

where A is the area of the section.

Equations (4) and (5) can be viewed as governing equations for plane strain displacements u and v where the terms multiplied by α are treated as body forces and similarly the α-terms in the boundary conditions (7) and (8) are treated as applied surface tractions. If we now use the reciprocal theorem with this field and the field for simple extension of a cylinder of cross section A, we are led to the result that

$$\iint_A \left(\frac{\partial u}{\partial \xi} + \frac{\partial v}{\partial \eta}\right) d\xi d\eta = -\frac{\lambda\alpha}{\lambda+\mu}\iint_A Dw\, d\xi d\eta. \quad (13)$$

Thus the expression for the axial load L can be written as

$$L = \mathrm{E}eA + E\alpha\iint_A \left(\eta\frac{\partial w}{\partial \xi} - \xi\frac{\partial w}{\partial \eta}\right) d\xi d\eta. \quad (14)$$

The integral in (14) can also be transformed by using the identity

$$\iint_A \left(\xi\frac{\partial w}{\partial \eta} - \eta\frac{\partial w}{\partial \xi}\right) d\xi d\eta = \iint_A \phi\left(\frac{\partial^2 w}{\partial \xi^2} + \frac{\partial^2 w}{\partial \eta^2}\right) d\xi d\eta - \int_C \phi\left(n_\xi\frac{\partial w}{\partial \xi} + n_\eta\frac{\partial w}{\partial \eta}\right) ds, \quad (15)$$

where $\phi(\xi,\eta)$ is the classical warping function for the section A and (n_ξ, n_η) is the unit normal to the bounding curve C of the region A in the ξ,η plane (n_ξ, n_η are proportional to $\partial F/\partial\xi, \partial F/\partial\eta$). The warping function $\phi(\xi,\eta)$ satisfies

$$\frac{\partial^2\phi}{\partial\xi^2} + \frac{\partial^2\phi}{\partial\eta^2} = 0 \text{ in } A, \quad (16)$$

$$\frac{\partial\phi}{\partial\xi}n_\xi + \frac{\partial\phi}{\partial\eta}n_\eta = \eta n_\xi - \xi n_\eta \text{ on } C,$$

and we also have the identity

$$\iint_A \left\{\frac{\partial G}{\partial\xi}\left(\frac{\partial\phi}{\partial\xi} - \eta\right) + \frac{\partial G}{\partial\eta}\left(\frac{\partial\phi}{\partial\eta} + \xi\right)\right\} d\xi d\eta = 0 \quad (17)$$

for any function $G(\xi,\eta)$ which is C_1 in A. With (6), (9), (15), (16), and the divergence theorem, equation (14) can be written as

$$L = \mathrm{E}eA + \alpha\mathrm{E}\left(\theta - \alpha\frac{\mathrm{E}e}{\mu}\right)(I-S) - \alpha^2\frac{\mathrm{E}}{\mu}\iint_A \left\{\lambda\left(\frac{\partial u}{\partial\xi} + \frac{\partial v}{\partial\eta}\right)D\phi + \mu\left[(-v+Du)\frac{\partial\phi}{\partial\xi} + (u+Dv)\frac{\partial\phi}{\partial\eta}\right]\right\} d\xi d\eta - \alpha^3\frac{\mathrm{E}}{\mu}(\lambda+2\mu)\iint_A D\phi Dw\, d\xi d\eta. \quad (18)$$

Here S is the geometrical torsional rigidity of the section and I is the moment of inertia of the section about the z-axis,

$$S = \iint_A \left\{\xi\left(\frac{\partial\phi}{\partial\eta} + \xi\right) - \eta\left(\frac{\partial\phi}{\partial\xi} - \eta\right)\right\} d\xi d\eta,$$

$$I = \iint_A (\xi^2+\eta^2) d\xi d\eta.$$

Similarly we can transform expression (12) for M to give

$$M = \mu\theta S + \alpha\mathrm{E}e(I-S) + \alpha\iint_A \left\{\lambda\left(\frac{\partial u}{\partial\xi} + \frac{\partial v}{\partial\eta}\right)D\phi + \mu\left(\xi\frac{\partial\phi}{\partial\xi} + \eta\frac{\partial\phi}{\partial\eta}\right)\left(\frac{\partial v}{\partial\xi} - \frac{\partial u}{\partial\eta}\right)\right\} d\xi d\eta + \alpha^2(\lambda+2\mu)\iint_A D\phi Dw\, d\xi d\eta. \quad (19)$$

The alternative expressions for L and M are useful when approximate solutions for small α are being found.

Because the state of strain is the same at each section, the axial force L on a section acts along the z-axis. The bending moments M_ξ, M_η on a section about the ξ and η axes are therefore zero. The moment M_ξ is given by

$$M_\xi = \iint_A \eta\left\{\mathrm{E}e + \lambda\left(\frac{\partial u}{\partial\xi} + \frac{\partial v}{\partial\eta}\right) + \alpha(\lambda+2\mu)Dw\right\} d\xi d\eta. \quad (20)$$

The term involving u and v in the integral can be transformed by treating u and v as the displacements for a plane strain problem and using the reciprocal theorem with the field for pure bending about the ξ-axis. With equations (4), (5), and boundary conditions (7) and (8) we then find that

$$M_\xi = \mathrm{E}\iint_A \eta(e+\alpha Dw) d\xi d\eta + \alpha\mu\nu\iint_A \left\{\frac{1}{2}(\xi^2-\eta^2)\left(\frac{\partial w}{\partial\xi} - \theta\eta\right) + \xi\eta\left(\frac{\partial w}{\partial\eta} + \theta\xi\right) + \alpha\left[\xi\eta(u+Dv) - \frac{1}{2}(\xi^2-\eta^2)(v-Du)\right]\right\} d\xi d\eta. \quad (21)$$

Similarly the bending moment M_η about the η-axis, given by

$$M_\eta = -\iint_A \xi\left\{\mathrm{E}e + \lambda\left(\frac{\partial u}{\partial\xi} + \frac{\partial v}{\partial\eta}\right) + \alpha(\lambda+2\mu)Dw\right\} d\xi d\eta \quad (22)$$

can be written as

$$M_\eta = -\mathrm{E}\iint_A \xi(e+\alpha Dw) d\xi d\eta - \alpha\mu\nu\iint_A \left\{\frac{1}{2}(\xi^2-\eta^2)\left(\frac{\partial w}{\partial\eta} + \theta\xi\right) - \xi\eta\left(\frac{\partial w}{\partial\xi} - \theta\eta\right) + \alpha\left[\xi\eta(v-Du) + \frac{1}{2}(\xi^2-\eta^2)(u+Dv)\right]\right\} d\xi d\eta. \quad (23)$$

Extension of a Rod With Small Initial Twist

We can consider extension and torsion separately and then use superposition for extension and torsion combined. In this section we determine the form of the solution for small pretwist α for the case of extension without torsion ($\theta = 0$).

Without loss of generality we can assume that the ξ-axis passes through the centroid of the cross section and we denote the distance of the centroid from the axis by R. When the centroids lie on the axis of pretwist ($R = 0$), as the pretwist α approaches zero the solution will approach that for simple extension of a cylinder, so that u, v, and w will tend to zero with α. However, as Okubo [2] has pointed out, when a spiral

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

rod with axis that does not pass through the centroids of the sections is pulled axially, the solution must reduce to extension combined with pure bending in the limit as the pretwist α becomes zero, because the axial load is not central to the section. Thus in this case the displacements u, v, and w are not small for small α.

We can also see this from the forms (20) and (22) for the bending moments on a section of the rod. Because we are taking the ξ-axis to pass through the centroid of the cross section, M_ξ will be zero when u, v, and w are zero, but in order to maintain M_η zero, the u, v, and w terms in (22) must make a $O(1)$ contribution to the integral in (22) as α tends to zero. With α set equal to zero in the equilibrium equations (4), (5), and the boundary equations (7) and (8), u and v become the solution of a plane strain problem with zero body force and zero surface tractions, and so represent a rigid displacement in the ξ, η plane. A rotation about the z-axis will not produce stresses in the pretwisted bar but leading terms in u and v that are constants will produce stresses which are of order α^2 times the constants, through coupling with the axial displacement w and through the terms multiplied by α in the governing equations. This suggests that for extension of the rod we take (u,v are even in α)

$$u = \frac{u_0}{\alpha^2} + O(1), \quad v = \frac{v_0}{\alpha^2} + O(1),$$

where u_0, v_0 are constants to be determined. Okubo omitted v_0 in his formulation [2], and we shall see that v_0 will be zero if the product of inertia of the section about the ξ,η axes is zero. The boundary condition (9) and equation (6) then suggest that

$$w = -\frac{u_0}{\alpha}\eta + \frac{v_0}{\alpha}\xi + O(\alpha).$$

With the assumption that the additional terms in the displacements can be represented as power series in α, we find from equations (4)–(9) with $\theta = 0$ that we have

$$\begin{aligned} u &= \frac{u_0}{\alpha^2} - \frac{1}{2}\nu u_0(\xi^2-\eta^2) - \nu v_0\xi\eta + u_2 \\ &\quad + \alpha^2 u'(\xi,\eta) + O(\alpha^4), \\ v &= \frac{v_0}{\alpha^2} - \nu u_0\xi\eta - \frac{1}{2}\nu v_0(\eta^2-\xi^2) + v_2 \\ &\quad + \alpha^2 v'(\xi,\eta) + O(\alpha^4), \\ w &= -\frac{u_0}{\alpha}\eta + \frac{v_0}{\alpha}\xi + \alpha(v_2\xi - u_2\eta) - \alpha\frac{\mathrm{E}e}{\mu}\phi(\xi,\eta) \\ &\quad - \frac{1}{3}\alpha u_0\eta^3 + \frac{1}{3}\alpha v_0\xi^3 + \alpha w'(\xi,\eta) + O(\alpha^3), \end{aligned} \tag{24}$$

aside from a rotation about the z-axis. The quantities u_2,v_2 are constants, and the function $w'(\xi,\eta)$ is harmonic in the cross section A and satisfies the condition

$$\begin{aligned} \frac{\partial w'}{\partial n} &= -u_0\left\{2\xi\eta n_\xi - (2\xi^2+\eta^2)n_\eta \right. \\ &\quad \left. + \nu\left[\xi\eta n_\xi - \frac{1}{2}(3\xi^2+\eta^2)n_\eta\right]\right\} \\ &\quad - v_0\left\{(2\eta^2+\xi^2)n_\xi - 2\xi\eta n_\eta + \nu\left[\frac{1}{2}(3\eta^2+\xi^2)n_\xi - \xi\eta n_\eta\right]\right\} \end{aligned} \tag{25}$$

on the boundary C of A.

With expressions (24), the displacements u,v become infinite as α tends to zero but following Okubo [2] we can add the rigid displacement

$$u_x = -\frac{u_0}{\alpha} + \frac{v_0}{\alpha}z, \quad u_y = -\frac{v_0}{\alpha^2} - \frac{u_0}{\alpha}z, \quad u_z = \frac{u_0}{\alpha}y - \frac{v_0}{\alpha}x.$$

As α tends to zero, the displacements now approach those corresponding to a rod in pure bending and extension.

The constants u_0,v_0 and u_2,v_2 in (24) are chosen so that M_ξ, M_η are zero to $O(1)$ and $O(\alpha^2)$, respectively. With expressions (21) and (23) for M_ξ and M_η and remembering that the ξ-axis passes through the centroid of the section, we find that u_0 and v_0 are given by

$$u_0 = -eRA/(R^2A + I_\eta - I_{\xi\eta}^2/I_\xi), \quad v_0 = -u_0 I_{\xi\eta}/I_\eta, \tag{26}$$

where the moments of inertia I_ξ, I_η, and $I_{\xi\eta}$ are defined by

$$I_\xi = \iint_A \eta^2 d\xi d\eta, \quad I_\eta = \iint_A (\xi - R)^2 d\xi d\eta,$$

$$I_{\xi\eta} = \iint_A \xi\eta d\xi d\eta = \iint_A (\xi - R)\eta d\xi d\eta.$$

We see that v_0 will be zero if $I_{\xi\eta}$ is zero.

When w' is known, u_2 and v_2 can be found by requiring expressions (21) and (23) for M_ξ and M_η to vanish to $O(\alpha^2)$. The value of the axial force L can then be found to $O(\alpha^2)$ from (14) or (18). From (14), (24), and (26) we have

$$L = \mathrm{E}eA(I_\eta - I_{\xi\eta}^2/I_\xi)/(R^2A + I_\eta - I_{\xi\eta}^2/I_\xi) + 0(\alpha^2). \tag{27}$$

The twisting moment M is $O(\alpha)$ and (19) and (24) give

$$\begin{aligned} M &= \alpha\mathrm{E}e(I-S) + \alpha\mathrm{E}\iint_A (u_0\xi + v_0\eta)D\phi d\xi d\eta \\ &\quad + 2\alpha\mu\nu\iint_A (v_0\xi - u_0\eta)\left(\xi\frac{\partial\phi}{\partial\xi} + \eta\frac{\partial\phi}{\partial\eta}\right)d\xi d\eta + O(\alpha^3), \end{aligned} \tag{28}$$

so that M can be found to $O(\alpha)$ when the torsion function $\phi(\xi,\eta)$ is known.

When the axis of pretwist passes through the centroids of the sections, the case $R = 0$, the constants u_0 and v_0 are zero so that w' can be taken to be zero. We therefore have in this case

$$\begin{aligned} u &= u_2 + \alpha^2 u'(\xi,\eta) + O(\alpha^4), \quad v = v_2 + \alpha^2 v'(\xi,\eta) + O(\alpha^4), \\ w &= \alpha(v_2\xi - u_2\eta) - \alpha\frac{\mathrm{E}}{\mu}e\phi(\xi,\eta) + O(\alpha^3). \end{aligned} \tag{29}$$

With expressions (21) and (23), u_2,v_2 are found to be given by

$$\begin{aligned} &u_2 I_\eta + v_2 I_{\xi\eta} - \frac{\mathrm{E}}{\mu}e\iint_A \xi D\phi d\xi d\eta \\ &\quad - e\nu\iint_A \left\{\frac{1}{2}(\xi^2-\eta^2)\frac{\partial\phi}{\partial\eta} - \xi\eta\frac{\partial\phi}{\partial\xi}\right\}d\xi d\eta = 0, \\ &u_2 I_{\xi\eta} + v_2 I_\xi - \frac{\mathrm{E}}{\mu}e\iint_A \eta D\phi d\xi d\eta \\ &\quad - e\nu\iint_A \left\{\frac{1}{2}(\xi^2-\eta^2)\frac{\partial\phi}{\partial\xi} + \xi\eta\frac{\partial\phi}{\partial\eta}\right\}d\xi d\eta = 0. \end{aligned} \tag{30}$$

For a section with two axes of symmetry, u_2 and v_2 are zero and

$$\begin{aligned} L &= \mathrm{E}eA\left\{1 - \alpha^2\frac{\mathrm{E}}{\mu}\frac{(I-S)}{A}\right\} + O(\alpha^4), \\ M &= \alpha\mathrm{E}e(I-S) + O(\alpha^3). \end{aligned} \tag{31}$$

Thus the axial stiffness is reduced by the pretwist in comparison with a cylindrical bar with the same section. A twisting moment is required to prevent the rod from unwinding as it is extended. We shall see later through an example that when the centroid of the section does not lie on the axis of pretwist, whether the rod will tend to unwind or to

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

twist further on extension depends on the shape of the cross section and the value of Poisson's ratio.

Torsion of a Rod With Small Initial Twist

For torsion with zero extension ($e=0$), the solution must make $O(\alpha)$ contributions to the integrals in (20) and (22) in order for M_ξ and M_η to vanish. This suggests that we have

$$u=\frac{u_1}{\alpha}+O(\alpha),\quad v=\frac{v_1}{\alpha}+O(\alpha)$$

for torsion of sections for which the integrals

$$\iint_A \eta D\phi d\xi d\eta,\quad \iint_A \xi D\phi d\xi d\eta$$

do not vanish, where u_1, v_1 are constants. Assuming expansions in α in equations (4)–(9) with $e=0$ leads to

$$\begin{aligned} u &= \frac{u_1}{\alpha}-\frac{1}{2}\alpha\nu u_1(\xi^2-\eta^2)-\alpha\nu v_1\xi\eta \\ &\quad +\alpha u_3+\alpha u'(\xi,\eta)+O(\alpha^3), \\ v &= \frac{v_1}{\alpha}-\alpha\nu u_1\xi\eta-\frac{1}{2}\alpha\nu v_1(\eta^2-\xi^2) \\ &\quad +\alpha v_3+\alpha v'(\xi,\eta)+O(\alpha^3), \\ w &= -u_1\eta+v_1\xi+\theta\phi(\xi,\eta)+\alpha^2 w'(\xi,\eta)+O(\alpha^4), \end{aligned} \tag{32}$$

apart from a rotation about the z-axis, where u_3, v_3 are constants. In order for M_ξ and M_η to vanish to $O(\alpha)$, u_1 and v_1 must satisfy

$$\begin{aligned} &u_1(R^2A+I_\eta)+v_1I_{\xi\eta}+\theta\iint_A \xi D\phi d\xi d\eta \\ &+\frac{\mu}{E}\nu\theta\iint_A\left\{\frac{1}{2}(\xi^2-\eta^2)\left(\frac{\partial\phi}{\partial\eta}+\xi\right)\right. \\ &\qquad \left.-\xi\eta\left(\frac{\partial\phi}{\partial\xi}-\eta\right)\right\}d\xi d\eta=0, \end{aligned} \tag{33}$$

$$\begin{aligned} &u_1I_{\xi\eta}+v_1I_\xi+\theta\iint_A \eta D\phi d\xi d\eta \\ &+\frac{\mu}{E}\nu\theta\iint_A\left\{\frac{1}{2}(\xi^2-\eta^2)\left(\frac{\partial\phi}{\partial\xi}-\eta\right)+\xi\eta\left(\frac{\partial\phi}{\partial\eta}+\xi\right)\right\}d\xi d\eta=0. \end{aligned}$$

If we add the rigid displacement

$$u_x=-\frac{u_1}{\alpha}+v_1z,\quad u_y=-\frac{v_1}{\alpha}-u_1z,\quad u_z=u_1y-v_1x,$$

then as α tends to zero, the displacements remain finite and approach those for pure torsion.

When u' and v' are known, the twisting moment M can be found to $O(\alpha^2)$ from (19). From (18), the axial force has the form

$$L=\alpha E\theta(I-S)+\alpha E\iint_A(u_1\xi+v_1\eta)d\xi d\eta+O(\alpha^3). \tag{34}$$

By using the relations (33) for u_1, v_1 and the values (26) for u_0, v_0 we can show that (34) and (28) satisfy reciprocity, that is the value of L given by (34) divided by θ is the same as the value of M given by (28) divided by e.

The equations for u' and v' are the equations of plane strain with body forces and surface tractions expressible in terms of the torsion function $\phi(\xi,\eta)$. The body force terms can be removed through a particular integral. One way is to set (compare [1])

$$u'=\theta U(\xi,\eta)-\theta(\eta\phi+\xi\psi-\Psi).\quad v'=\theta V(\xi,\eta), \tag{35}$$

where the harmonic functions ψ and Ψ are defined through

$$\zeta=\xi+i\eta,\quad \phi+i\psi=f'(\zeta),\quad \Phi+i\Psi=f(\zeta).$$

The functions U,V satisfy the equilibrium equations of plane strain with no body force (set $\alpha=0$ in (4) and (5)) together with the boundary conditions

$$\begin{aligned} &\left\{(\lambda+2\mu)\frac{\partial U}{\partial\xi}+\lambda\frac{\partial V}{\partial\eta}\right\}n_\xi+\mu\left(\frac{\partial U}{\partial\eta}+\frac{\partial V}{\partial\xi}\right)n_\eta \\ &\quad=2\mu\left\{\left(\eta^2-\xi\frac{\partial\phi}{\partial\eta}\right)n_\xi+\left(\xi\frac{\partial\phi}{\partial\xi}-\xi\eta\right)n_\eta\right\}, \\ &\left\{(\lambda+2\mu)\frac{\partial V}{\partial\eta}+\lambda\frac{\partial U}{\partial\xi}\right\}n_\eta+\mu\left(\frac{\partial U}{\partial\eta}+\frac{\partial V}{\partial\xi}\right)n_\xi=0 \end{aligned} \tag{36}$$

on the boundary C of A.

When the axis of pretwist passes through the centroids of the sections, the constants u_1, v_1 are not necessarily zero, but they will vanish if in addition the section has two axes of symmetry. In this case we have

$$M=\mu\theta S+O(\alpha^2),\quad L=\alpha E\theta(I-S)+O(\alpha^3), \tag{37}$$

and with our formulation, u' and v' must be found in order to determine M to $O(\alpha^2)$. On twisting, the rod tends to contract or extend according as the twist increases or decreases the pretwist.

Combined Extension and Torsion

Results for combined extension and torsion can be obtained by superposing the results of the two preceding sections. In this section we suppose that the axis of pretwist passes through the centroid of the cross section and that the section has two axes of symmetry. From (31) and (37) we then have

$$\begin{aligned} L&=EeA\left\{1-\alpha^2\frac{E}{\mu}\frac{(I-S)}{A}\right\}+\alpha E\theta(I-S)+\theta O(\alpha^3)+eO(\alpha^4), \\ M&=\mu\theta S+\alpha Ee(I-S)+\theta O(\alpha^2)+eO(\alpha^3). \end{aligned} \tag{38}$$

If we take

$$\theta=-\alpha Ee(I-S)/\mu S, \tag{39}$$

then M will be zero to $O(\alpha^2)$ and we are approximating the case of extension with zero twisting moment. The axial load is then

$$L=EeA\left\{1-\alpha^2\frac{E}{\mu}\frac{I}{S}\frac{(I-S)}{A}\right\}+O(\alpha^4). \tag{39}$$

Comparison with (31) shows that the axial stiffness is less with zero moment than with zero twist because $S\le I$. The more slender the section, the smaller S is in comparison to I and the greater the reduction in the axial stiffness and the larger the anount of unwinding needed in order to reduce the moment to zero.

If we take

$$e=-\alpha\theta(I-S)/A, \tag{40}$$

then the axial load will be zero to $O(\alpha^2)$ and the twisting moment will be reduced by an amount

$$\alpha^2E\theta(I-S)^2/A+O(\alpha^4) \tag{41}$$

in comparison to twisting with no extension.

Elliptical Cross Section

For an elliptical cross section with semiaxes a, b ($a>b$) and

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

centroid on the z-axis, we can take the boundary C of A to be given by

$$\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} = 1.$$

For the elliptical cross section we have

$$\phi = -\frac{(a^2-b^2)}{(a^2+b^2)}\xi\eta, \quad I = \frac{\pi}{4}ab(a^2+b^2), \quad S = \frac{\pi a^3 b^3}{(a^2+b^2)}.$$

For torsion of the rod, the constants u_1, v_1 of equations (32) will vanish from symmetry, and the constants u_3, v_3 can be taken to be zero. The functions U, V introduced in (35) satisfy the plane strain equilibrium equations and the boundary conditions (36), and we find that u and v are given by

$$\frac{u}{\alpha\theta}\frac{(1+\rho)}{(1-\rho)}(3+2\rho+3\rho^2) = -(1+3\rho)[\rho+\nu(3+\rho)]\xi^3/3 - (1+3\rho)[3-\nu(3+\rho)]\xi\eta^2 + \rho[1+3\rho+2\nu(1-\rho)]a^2\xi + O(\alpha^2), \tag{43}$$

$$\frac{v}{\alpha\theta}\frac{(1+\rho)}{(1-\rho)}(3+2\rho+3\rho^2) = (3+\rho)[1+\nu(3\rho+1)]\eta^3/3 + (3+\rho)[3\rho-\nu(3\rho+1)]\xi^2\eta - \rho[\rho+3-2\nu(1-\rho)]a^2\eta + O(\alpha^2),$$

where $\rho = b^2/a^2$.

Although the displacements are known only to $O(\alpha)$, the twisting moment M can now be found to $O(\alpha^2)$ by substitution in (19) and we find that

$$M/\mu\theta S = 1 + \alpha^2 a^2(1-\rho)^2\{9-6\rho-22\rho^2-6\rho^3+9\rho^4 + \nu(9+24\rho+62\rho^2+24\rho^3+9\rho^4)\}/12\rho(1+\rho)(3+2\rho+3\rho^2) + O(\alpha^4). \tag{44}$$

The contribution of the α^2 term is positive for positive ν if $\rho < 0.511$ or $b/a < 0.715$. If $\nu > 1/8$ then the α^2 term is positive for all b/a.

The axial force L can be found to $O(\alpha^3)$ from (18),

$$L/\theta a E A = \frac{\alpha a}{4}\frac{(1-\rho)^2}{(1+\rho)}\left\{1-\alpha^2 a^2[9+2\rho-6\rho^2+2\rho^3+9\rho^4 + 2\nu(3+4\rho+18\rho^2+4\rho^3+3\rho^4)]/3(1+\rho)(3+2\rho+3\rho^2)\right\} + O(\alpha^5). \tag{45}$$

As shown in the preceding section, if we add an extension

$$e = -\alpha\theta(I-S)/A = -\alpha\theta a^2(1-\rho)^2/4(1+\rho) \tag{46}$$

then L will be $O(\alpha^3)$. The extension reduces the twisting moment by the amount (42) so that the moment for torsion with no axial load is given by

$$M/\mu\theta S = 1 + \alpha^2 a^2(1-\rho)^2[9-50\rho^2+9\rho^4+\nu(9+60\rho + 118\rho^2+60\rho^3+9\rho^4)]/24\rho(1+\rho)(3+2\rho+3\rho^2) + O(\alpha^4). \tag{47}$$

The moment is less with the constraint of no extension removed. The α^2 term is again positive if $\nu > 1/8$ and is positive for positive ν if $b/a < 0.657$.

The increase in torsional rigidity because of the pretwist is appreciable for elongated sections for which ρ is small. Thus for $a/b = 10$, $\nu = 0.3$, and $\alpha a = 0.15$, the second order forms (44) and (47) indicate 71 percent and 36 percent increases in the torsional stiffness for the cases of zero extension and zero axial load, respectively. For small ρ we have for torsion with no extension

$$M/\mu\theta S = 1 + (1+\nu)\alpha^2 a^2/4\rho + O(\alpha^2 a^2)$$

and for zero axial load

$$M/\mu\theta S = 1 + (1+\nu)\alpha^2 a^2/8\rho + O(\alpha^2 a^2). \tag{48}$$

The latter expression agrees with that given by Chen Chu [4], who used an approximate analysis for torsion of twisted bars of slender section under no axial load. Chu found good agreement between theory and experiment for torsion of narrow steel strips of constant thickness for amounts of pretwist up to $\alpha a = 0.16$, where a is half the width of the strip.

The strains can be found to $O(\alpha)$ by substitution in (10). The axial stress σ_{zz} is found to be given by

$$\frac{\sigma_{zz}}{\alpha\mu\theta}\frac{(1+\rho)}{(1-\rho)}(3+2\rho+3\rho^2) = 2(\xi^2-\eta^2)\{3+2\rho+3\rho^2 + \nu(1+3\rho)(3+\rho)\} - 4a^2\nu\rho(1-\rho) + O(\alpha^2) \tag{49}$$

for torsion with no extension. The axial stress in the outer fibers is tensile and the axial stress in the center fibers is compressive or vice versa when θ has the same sign as α or the opposite. For torsion with no axial load we add the axial stress Ee with e given by (46) and the corresponding form to (49) is then

$$\frac{1}{2}\left\{4(\xi^2-\eta^2) - a^2(1-\rho)\right\} \times \left\{3+2\rho+3\rho^2+\nu(1+3\rho)(3+\rho)\right\}. \tag{50}$$

In both cases the maximum axial stress occurs at the ends of the major axis. For small ρ

$$(\sigma_{zz}/\mu\theta a)_{(a,0)} = \begin{cases} 2\alpha a(1+\nu) + O(\alpha b) & \text{for } e = 0, \\ \frac{3}{2}\alpha a(1+\nu) + O(\alpha b) & \text{for } L = 0. \end{cases}$$

We see that for slender sections, the axial stress can be comparable to the value

$$2\mu\theta a\rho^{1/2}/(1+\rho) + O(\alpha^2)$$

for the maximum shear stress on a plane $z = \text{constant}$.

The torsion of a twisted bar of elliptical section with no axial load was treated by Okubo [1] and he obtained values for the coefficients entering into the forms for the displacements of $O(\alpha)$ for the particular case $a/b = 2$ and $\nu = 0.2568$. His value for the extension is in agreement with equation (46) and the coefficients lead to an expression for σ_{zz} which is in agreement with the form (50).

For extension of the rod, the constants u_0, v_0 and u_2, v_2 appearing in (24) are zero and w' can be taken to be zero. Thus w is known to $O(\alpha)$ and u and v can be found to $O(\alpha^2)$ from the first two equilibrium equations (4) and (5) and the first two boundary conditions (7) and (8) with $\theta = 0$. We then have

$$\frac{u}{\alpha^2 e}\frac{(1+\rho)(3+2\rho+3\rho^2)}{(1+\nu)(1-\rho)} = [1+2\rho+5\rho^2+2\nu(1+6\rho+\rho^2)]\xi^3/3 + [1+10\rho-3\rho^2-2\nu(1+6\rho+\rho^2)]\xi\eta^2 - [1+2\rho+5\rho^2-\nu(1-\rho)^3]a^2\xi + O(\alpha^2),$$

$$\frac{v}{\alpha^2 e}\frac{(1+\rho)(3+2\rho+3\rho^2)}{(1+\nu)(1-\rho)} = -[5+2\rho+\rho^2+2\nu(1+6\rho+\rho^2)]\eta^3/3 \tag{51}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+[3-10\rho-\rho^2+2\nu(1+6\rho+\rho^2)]\xi^2\eta$$

$$+[\rho(5+2\rho+\rho^2)+\nu(1-\rho)^3]a^2\eta+O(\alpha^2),$$

$$\frac{w}{\alpha e}\frac{(1+\rho)}{(1+\nu)(1-\rho)}=2\xi\eta+O(\alpha^2).$$

The axial load is given by (31) to $O(\alpha^2)$ and we have

$$L/\mathrm{E}eA=1-\alpha^2a^2(1+\nu)(1-\rho)^2/2(1+\rho)+O(\alpha^4). \tag{52}$$

If desired, L can be determined to $O(\alpha^4)$ by substituting (51) into (18). The moment M can be found to $O(\alpha^3)$ by substituting (51) into (19) and it is found that $M/ea\mathrm{E}A$ is given by the right-hand side of (45), in agreement with reciprocity.

According to (52), the axial force for extension with no torsion does not change appreciably with small amounts of pretwist. However if we add torsion of amount given by (39), that is

$$\theta=-\alpha e(1+\nu)(1-\rho)^2/2\rho, \tag{53}$$

in order to reduce the moment to $O(\alpha^3)$, then the axial load is from (39)

$$L/\mathrm{E}eA=1-\alpha^2a^2(1+\nu)(1-\rho)^2(1+\rho)/8\rho+O(\alpha^4). \tag{54}$$

We see that for elongated sections for which ρ is small, the axial load for extension with no moment can be considerably less than that for extension with no torsion even for small amounts of pretwist.

The axial stress for extension with no torsion is given by

$$\sigma_{zz}/\mathrm{E}e=1-\alpha^2(1-\rho)\{2(\xi^2-\eta^2)[3+2\rho+3\rho^2+\nu(1+6\rho+\rho^2)]$$
$$+a^2\nu(1-\rho)^3\}/(1+\rho)(3+2\rho+3\rho^2)+O(\alpha^4). \tag{55}$$

We can use (49) to find the axial stress induced by torsion of amount (54), and adding this stress to the value given by (55) we obtain

$$\sigma_{zz}/\mathrm{E}e=1-\alpha^2(1+\nu)(1-\rho^2)(\xi^2-\eta^2)/2\rho+O(\alpha^4) \tag{56}$$

for extension with no twisting moment. The shear stresses $\sigma_{\xi z},\sigma_{\eta z}$ for this case are

$$\sigma_{\xi z}=\alpha\mu e\eta(1+\nu)(1-\rho)/\rho+O(\alpha^3),$$
$$\sigma_{\eta z}=\alpha\mu e\xi(1+\nu)(1-\rho)+O(\alpha^3). \tag{57}$$

We see that the corrections to the axial stress and the shear stress $\sigma_{\xi z}$ become appreciable when b/a is comparable to αa.

Okubo [1] considered extension of a twisted elliptical bar with no moment and gave results to $O(\alpha^2)$ for the case $a/b=2$ and $\nu=0.2568$. In this case (54) gives

$$L/\mathrm{E}eA=1-1.7674\alpha^2b^2+O(\alpha^4),$$

which agress with Okubo except for the last decimal place in the coefficient of α^2b^2. His value for the torsion required to nullify the twisting moment agrees with (53) and the expressions for the axial stress and the shear stresses $\sigma_{\xi z}$ and $\sigma_{\eta z}$ are in agreement with (56) and (57).

Off-Centered Circular Section

The simplest section to consider which does not have its centroid on the axis of pretwist is the off-centered circular section with boundary C given by

$$(\xi-R)^2+\eta^2=a^2,$$

where a is the radius of the section. We then have

$$\phi=-R\eta,\quad I=\pi a^2(R^2+a^2/2),\quad S=\pi a^4/2.$$

For torsion with no extension, v_1 and v_3 of (32) will be zero from symmetry and from (33) we obtain

$$u_1=-\theta R\{a^2+4R^2+2\nu(a^2+2R^2)\}/(1+\nu)(a^2+4R^2). \tag{58}$$

The displacements are found to $O(\alpha)$ by solving the associated plane strain problem for u', v' and we arrive at

$$u=\frac{u_1}{\alpha}-\frac{1}{2}\alpha\nu u_1(\xi^2-\eta^2)+\alpha u_3$$
$$-\alpha\theta R[\nu\xi'^2+(1-\nu)\eta^2+\nu R\xi']+O(\alpha^3),$$
$$v=-\alpha\nu u_1\xi\eta+\alpha\theta R[(1-2\nu)\xi'\eta-\nu R\eta]+O(\alpha^3), \tag{59}$$
$$w=-\theta R\eta-u_1\eta+O(\alpha^2),$$

where ξ' is used to denote ξ-R. The constant u_3 is determined by the condition that M_η vanish to $O(\alpha^3)$, and with the form (23) for M_η we need to know w to $O(\alpha^2)$ in order to determine u_3. However, at this stage we can determine M to $O(\alpha^2)$ from (19) and we have

$$M/\mu\theta S=1-\alpha^2R^2\{(3+\nu)(1+4r)-8\nu^2r\}/2(1+\nu)(1+4r)$$
$$+O(\alpha^4), \tag{60}$$

where r denotes R^2/a^2.

The axial load L is given by (34) to $O(\alpha)$,

$$L/\mu\theta R\pi a^2=-2\alpha R\nu/(1+4r)+O(\alpha^3). \tag{61}$$

We see that for small amounts of pretwist, with the length held constant, L is compressive or tensile according, as the torsion increases or decreases the pretwist.

With the values (59) for u and v, w can be found to $O(\alpha^2)$ from the equilibrium equation (6) and the boundary equation (9), and we find that

$$w=-(u_1+\theta R)\eta-\alpha^2u_3\eta$$
$$+\alpha^2u_1\left\{-\frac{1}{4}(\xi'^2\eta+\eta^3)+(1+\nu)R\xi'\eta\right.$$
$$\left.+\left[\frac{1}{4}(3+2\nu)a^2+\frac{1}{2}(4+3\nu)R^2\right]\eta\right\}$$
$$+\alpha^2\theta R\left\{-\frac{3}{8}(\xi'^2\eta+\eta)+\frac{1}{2}(1+4\nu)R\xi'\eta\right.$$
$$\left.+\left[\frac{1}{8}(9+8\nu)a^2+(2+\nu)R^2\right]\eta\right\}+O(\alpha^4). \tag{62}$$

For M_η to vanish to $O(\alpha^3)$, we need

$$u_3=\theta Ra^2\{(1+4r)(7+12r)+\nu(15+116r+240r^2-192r^3)$$
$$+6\nu^2(1+6r+24r^2-64r^3)$$
$$-16\nu^3r(1+3r+12r^2)\}/24(1+\nu)^2(1+4r)^2. \tag{63}$$

If desired, L can now be found to $O(\alpha^3)$ by substitution in (18).

For extension of the rod, v_0 and v_2 in (24) are zero from symmetry and from (26),

$$u_0=-4eR/(a^2+4R^2). \tag{64}$$

The harmonic function w' satisfying the boundary condition (25) is

$$w'=-u_0\left\{\frac{1}{4}\left(\xi^2\eta-\frac{1}{3}\eta^3\right)\right.$$
$$\left.-\frac{1}{2}(3+2\nu)R\xi\eta-\frac{1}{4}(3+2\nu)(a^2+R^2)\eta\right\} \tag{65}$$

apart from a constant. The constant u_2 can now be determined by requiring that the form (23) for M_η vanish to $O(\alpha^2)$,

$$u_2=-eR\{1+12r+6\nu r(3-4r)$$
$$-2\nu^2(1+3r+12r^2)\}/3(1+\nu)(1+4r)^2. \tag{66}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The axial force L can now be found to $O(\alpha^2)$ from (18) and we have

$$L/\mathrm{Ee}A = \frac{1}{(1+4r)} - \alpha^2 R^2 \frac{(1+12r+24\nu r-2\nu^2)}{3(1+\nu)(1+4r)^2} + O(\alpha^4). \quad (67)$$

From (28) the twisting moment induced by the extension is given by

$$M/\mu eR\pi a^2 = -2\alpha R\nu/(1+4r) + O(\alpha^3), \quad (68)$$

in agreement with the axial load (61) induced by torsion.

The moment and axial load for combined extension and torsion can be derived by superposition. For torsion with no axial load, if we add an extension e given by

$$e = \alpha\theta R^2\nu/(1+\nu), \quad (69)$$

the axial load will be $O(\alpha^3)$ and the moment is then given by

$$M/\mu\theta S = 1 - \alpha^2 R^2(3+\nu)/2(1+\nu) + O(\alpha^4). \quad (70)$$

For extension with no twisting moment, M will be $O(\alpha^3)$ if we add torsion of amount

$$\theta = 4\alpha er\nu/(1+4r), \quad (71)$$

and then the axial load is given by

$$L/\mathrm{Ee}A = \frac{1}{(1+4r)} - \alpha^2 R^2 \frac{\{1+12r+24\nu r-2\nu^2(1-6r)\}}{3(1+\nu)(1+4r)^2} + 0(\alpha^4). \quad (72)$$

For a fixed value of αa, the second-order terms in the expressions (70) and (72) for the moment for torsion and the axial load for extension become more important as R/a increases. However the present theory is only valid for R/a not too large. At large values of R/a geometry changes must be taken into account even in a small strain theory.

The case of extension with no moment for an off-centered circular section was considered by Okubo [2], and in our notation he found that to $O(\alpha)$ the induced torsion is $\theta = -\alpha R u_0 \nu$ which is in agreement with (71). Okubo treated the particular case $\alpha a = 0.25$ and $\nu = 0.3$ for values of R/a up to 1.5 and gave results on the axial and shearing stresses.

For the off-centered elliptical section with boundary given by

$$(\xi - R)^2/a^2 + \eta^2/b^2 = 1,$$

the classical warping function is

$$\phi = -R\eta - (a^2 - b^2)(\xi - R)\eta/(a^2 + b^2).$$

For extension of the rod, v_0 is zero and u_0 is still given by (64). From (28) the induced twisting moment is given by

$$M/\mathrm{Ee}aA = \frac{\alpha a\{(1-\rho)^2 - 4r[1-\rho(1-\nu)/(1+\nu)]\}}{4(1+\rho)(1+4r)} + O(\alpha^3), \quad (73)$$

where, as before, $r = R^2/a^2$ and $\rho = b^2/a^2$. For sections close to a circle, equation (73) shows that for small amounts of pretwist, the pretwist tends to increase or decrease when the rod is pulled in tension or compressed, respectively. When $R/a < 1/2$, rods with elongated sections (ρ small or large enough) tend to unwind on extension. For $R/a > 1/2$, the rod will tend to unwind or not on extension according as ρ is greater or less than a critical value that lies between $(1+3\nu)/(1+\nu)$ and $(1+\nu)/(1-\nu)$. Similar remarks apply, through reciprocity, to the tendency of the bar to extend or shorten when twisted.

The results of Reissner and Wan [10] for helicoidal shells show that a slightly twisted strip with an off-centered narrow rectangular section will tend to unwind or twist further on extension accordingly as the ratio of the distance of the center of the section from the axis to the half width of the section is less than or greater than $1/\sqrt{3}$, which is close to the critical value $R/a = 1/2$ derived here for the off-centered elliptical section highly elongated in the radial direction ($\rho \to 0$).

References

1 Okubo, H., "The Torsion and Stretching of Spiral Rods I," *Quarterly of Applied Mathematics*, Vol. 9, 1951, pp. 263–272.

2 Okubo, H., "The Torsion and Stretching of Spiral Rods II," *Quarterly of Applied Mathematics*, Vol. 11, 1954, pp. 488–495.

3 Okubo, H., "The Torsion of Spiral Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 20, 1953, pp. 273–278.

4 Chu, C., "The Effect of Initial Twist on the Torsional Rigidity of Thin Prismatical Bars and Tubular Members," *Proceedings of the First U.S. National Congress on Applied Mechanics*, Chicago, Ill., June 1951, pp. 265–269.

5 Rosen, A., "The Effect of Initial Twist on the Torsional Rigidity of Beams—Another Point of View," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 389–392.

6 Hodges, D. H., "Torsion of Pretwisted Beams due to Axial Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 393–397.

7 Rosen, A., Discussion of Reference 6 and Author's Closure, ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 679–681.

8 Washizu, K., "Some Considerations on a Naturally Curved and Twisted Slender Beam," *Journal of Mathematics and Physics*, Vol. 43, 1964, pp. 111–116.

9 Knowles, J. K., and Reissner, E., "Torsion and Extension of Helicoidal Shells," *Quarterly of Applied Mathematics*, Vol. 17, 1960, pp. 409–422.

10 Reissner, E., and Wan, F. Y. M., "On Axial Extension and Torsion of Helicoidal Shells," *Journal of Mathematics and Physics*, Vol. 47, 1968, pp. 1–31.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

E. E. Gdoutos
Professor of Applied Mechanics.
Mem. ASME

M. Kattis
Senior Student.

School of Engineering,
Democritus University of Thrace,
Xanthi, Greece

Stress Function Interface and Boundary Conditions in Anisotropic Materials

The stress and displacement continuity conditions for interfaces between two different anisotropic media were formulated in terms of the Airy stress functions of the media. It was shown that such formulation greatly facilitates the solution of the problems of composite anisotropic materials by the Airy stress function approach. Two examples were given to demonstrate the potentiality of the method.

Introduction

The Airy stress function method of two-dimensional elasticity has extensively been used for the solution of beam problems for various geometrical configurations and loading conditions. A survey of such problems for one-layer isotropic beams was provided by Timoshenko and Goodier [1]. Lekhnitskii [2] used the same approach for solving problems concerning anisotropic beams and plates.

For obtaining closed-form elasticity solutions in multilayered beams by the Airy stress function method, the Airy function should be determined in each separate layer using the appropriate boundary and continuity conditions. These conditions express the fact that the stresses and displacements should satisfy certain prescribed requirements along the boundaries or the interfaces of the composite body. Determination of the stresses from the Airy stress function needs two successive differentiations and it is a relatively simple process, whereas for the corresponding calculation of the displacements, one integration is needed. During the latter process, integration constants are entered that complicate the solution of the problem. Therefore, it seems desirable to express the displacement continuity conditions in terms of the Airy stress function.

Zienkiewicz and Gerstner [3] were the first to obtain the stress and displacement continuity conditions for an interface separating isotropic materials with different elastic properties. Furthermore, they expressed these conditions in terms of their finite-difference approximations and gave an example illustrating the method. Schile [4] presented an elementary solution of the plane-stress equation of nonhomogeneous isotropic elasticity for the case when the elastic constants are functions of one variable. To illustrate the method he gave the solution of a sandwich two-material cantilever subjected to an end force and moment. The same author [5] considered the problem of a nonhomogeneous prismatic bar of arbitrary cross section subjected to a terminal force and moment and derived a partial differential equation for the stress function. For a small deviation from homogeneity, he obtained the solution as a perturbation expansion near the homogeneous solution and gave an analytical solution for an elliptical bar with a particular type of inhomogeneity. Furthermore, Schile [6] presented a general formulation for the plane elastostatic laminated body problem. He showed that the problem involves the determination of a single piecewise continuous stress function which was expressed in terms of the material property functions and three continuous stress functions which have discontinuous normal derivatives at material interfaces. Also, he studied the problem of laminated plates made of isotropic plies by considering the laminated plate as a material continuum characterized by material property functions with finite jumps along material interfaces [7, 8].

The bending problem of laminated plates and shells consisting of anisotropic materials was studied by Reissner and Stavsky [9] and Dong et al. [10]. In [9], the case of plates consisting of two orthotropic sheets of equal thickness which are laminated in such a way that the axes of elastic symmetry enclose an angle $+\theta$ with the coordinate axes in one sheet and an angle $-\theta$ in the other was studied. In [10] a small deflection theory governing the elastostatic extension of flexure of thin laminated anisotropic plates and shells was formulated. Pagano [11] studied the limitations of the laminated theory introduced in [9, 10] by comparing solutions of several specific boundary value problems in this theory to the corresponding solutions of the theory of elasticity. He found that the laminated theory gives adequate results for relatively large span-to-depth ratios.

Gerstner [12] presented a closed-form solution for a composite cantilever consisting of three isotropic layers subjected to an end force using the Airy stress function method. He found that the results are identical to those obtained from the usual engineering solution. Sierakowski and Ebcioglu [13] using the results of reference [4] obtained a solution for the case of a multilayered cantilever. As a special case they considered the symmetrical cantilever consisting of a

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division June, 1981; final revision, January, 1982. Paper No. 82-WA/APM-4.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

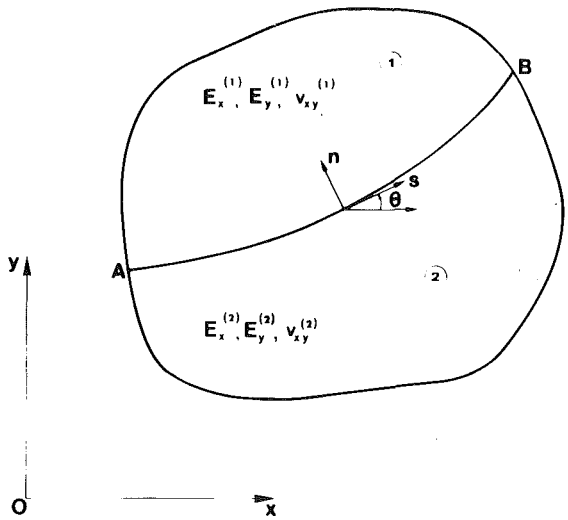


Fig. 1 Two dissimilar orthotropic materials bonded along the interface AB

soft core and hard faces and subjected to an end force. Whitney et al. [14] modified the Pagano plate theory [11] for application to beams. Thus, they derived equations for the general case of symmetrically laminated beams by considering the beam as a special case of the laminated plate. By comparing the results of their theory to flexure data on graphite/epoxy angle-ply and quasi-isotropic laminates they found that the theoretical results are in good agreement with the experiments.

In the present paper, the stress and displacement continuity conditions for interfaces between bonded anisotropic media with different material constants are formulated in terms of the corresponding Airy stress functions of the media. Such a formulation greatly facilitates the solution of laminated anisotropic beam problems. The usefulness and potentiality of the obtained conditions is demonstrated by two examples, namely, the two-material composite cantilever loaded by an end-concentrated load and moment, and the two-material composite simply supported beam loaded by a uniform load along its upper boundary.

Displacement Continuity Conditions

Consider two media made of rectilinearly orthotropic materials with different elastic constants bonded together along the line AB (Fig. 1). The materials are described by their elastic constants E_x^i, E_y^i, ν_{xy}^i, ν_{yx}^i $(i=1, 2)$ $(E_x^i/E_y^i=\nu_{xy}^i/\nu_{yx}^i)$ referred to their principal orthotropic directions x, y, which are assumed to be common for both media. On the interface AB no slippage is possible and therefore the displacements in both media can differ only by the component of a rigid body displacement. Thus, if u and v denote the displacements in the x and y directions the following conditions should be satisfied on the interface:

$$u_1 = u_2 + \alpha y + \beta$$
$$v_1 = v_2 - \alpha x + \delta \tag{1}$$

where the terms $(\alpha y+\beta)$ and $(-\alpha x+\delta)$ denote the rigid-body displacements with α, β, δ constants.

To eliminate the constants α, β, δ of the rigid-body displacement we differentiate relations (1) with respect to the variable s along the interface. Thus, the following relations are obtained:

$$\frac{\partial u_1}{\partial s}\cos\theta+\frac{\partial v_1}{\partial s}\sin\theta=\frac{\partial u_2}{\partial s}\cos\theta+\frac{\partial v_2}{\partial s}\sin\theta$$

$$-\frac{\partial^2 u_1}{\partial s^2}\sin\theta+\frac{\partial^2 v_1}{\partial s^2}\cos\theta=-\frac{\partial^2 u_2}{\partial s^2}\sin\theta+\frac{\partial^2 v_2}{\partial s^2}\cos\theta \tag{2}$$

where θ is the angle made by the tangent to the interface and the x-axis.

Relations (2) will now be expressed in terms of displacements u_i' and v_i' $(i = 1, 2)$ along and normal to the interface. Thus, by taking into account relations:

$$u = u'\cos\theta - v'\sin\theta$$
$$v = u'\sin\theta + v'\cos\theta \tag{3}$$

we obtain:

$$\frac{\partial u}{\partial s}\cos\theta+\frac{\partial v}{\partial s}\sin\theta=\frac{\partial u'}{\partial s}-\frac{v'}{\rho}$$

$$-\frac{\partial^2 u}{\partial s^2}\sin\theta+\frac{\partial^2 v}{\partial s^2}\cos\theta=\frac{\partial^2 v'}{\partial s^2}+\frac{1}{\rho}\frac{\partial u'}{\partial s}+u'\frac{\partial}{\partial s}\left(\frac{1}{\rho}\right) \tag{4}$$

where $\rho = (ds/d\theta)$ represents the radius of curvature of the interface. Thus, relations (2) yield:

$$\frac{\partial u_1'}{\partial s}-\frac{v_1'}{\rho}=\frac{\partial u_2'}{\partial s}-\frac{v_2'}{\rho}$$

$$\frac{\partial^2 v_1'}{\partial s^2}+\frac{1}{\rho}\frac{\partial u_1'}{\partial s}+u_1'\frac{\partial}{\partial s}\left(\frac{1}{\rho}\right)=\frac{\partial^2 v_2'}{\partial s^2}+\frac{1}{\rho}\frac{\partial u_2'}{\partial s}+u_2'\frac{\partial}{\partial s}\left(\frac{1}{\rho}\right) \tag{6}$$

The stresses σ_x, σ_y, τ_{xy} are expressed in each of the two media in terms of the corresponding Airy stress function ϕ by the following relations:

$$\sigma_x=\frac{\partial^2\phi}{\partial y^2}\qquad \sigma_y=\frac{\partial^2\phi}{\partial x^2}\qquad \tau_{xy}=-\frac{\partial^2\phi}{\partial x\partial y} \tag{7}$$

The strains are given according to Hooke's stress-strain law for generalized plane stress conditions by the following relations:

$$\epsilon_x = \frac{\partial u}{\partial x}=\frac{\sigma_x}{E_x}-\frac{\nu_{yx}}{E_y}\sigma_y$$
$$\epsilon_y = \frac{\partial v}{\partial y}=\frac{\sigma_y}{E_y}-\frac{\nu_{xy}}{E_x}\sigma_x$$
$$\gamma_{xy} = \frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}=\frac{\tau_{xy}}{G_{xy}} \tag{8}$$

Introducing the notations:

$$A=(E_xE_y)^{-(1/2)}\qquad B=(\nu_{xy}\nu_{yx})^{1/2}\qquad G_{xy}=\frac{1}{2A(\varkappa+B)}$$

$$\frac{E_x}{E_y}=\frac{\nu_{xy}}{\nu_{yx}}=c^2 \tag{9}$$

and taking into account relations (7) we obtain:

$$\frac{\partial u}{\partial x}=A\left(\frac{1}{c}\frac{\partial^2\phi}{\partial y^2}-B\frac{\partial^2\phi}{\partial x^2}\right)$$

$$\frac{\partial v}{\partial y}=A\left(c\frac{\partial^2\phi}{\partial x^2}-B\frac{\partial^2\phi}{\partial y^2}\right)$$

$$\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}=-2A(\kappa+B)\frac{\partial^2\phi}{\partial x\partial y} \tag{10}$$

For the case of plane strain conditions constants A, B, and c take the values:

$$A = (E_x'E_y')^{-(1/2)}\qquad B=(\nu_{xy}'\nu_{yx}')^{1/2} \tag{11}$$
$$c^2 = \frac{E_x'}{E_y'}=\frac{\nu_{xy}'}{\nu_{yx}'}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with:

$$E_x' = \frac{E_x}{1-\nu_{zx}\nu_{xz}} \qquad E_y' = \frac{E_y}{1-\nu_{zy}\nu_{yz}}$$

$$\nu_{xy}' = \frac{\nu_{xy}+\nu_{zy}\nu_{xz}}{1-\nu_{zx}\nu_{xz}} \qquad \nu_{yx}' = \frac{\nu_{yx}+\nu_{yz}\nu_{zx}}{1-\nu_{zy}\nu_{yz}} \tag{12}$$

where the constants ν_{ij}' $(i, j = x, y, z)$ are analogous to the Poisson's ratios ν_{xy}, ν_{yx} entered into relations (8).

We now have from the first of relations (4) that:

$$\frac{\partial u'}{\partial s} - \frac{v'}{\rho} = \frac{\partial u}{\partial x}\cos^2\theta + \frac{\partial v}{\partial y}\sin^2\theta + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)\sin\theta\cos\theta \tag{13}$$

and taking into account relations (10) we obtain:

$$\frac{\partial u'}{\partial s} - \frac{v'}{\rho} = \frac{A}{c}\left[R - Bc\left(\frac{\partial^2\phi}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi}{\partial n}\right)\right] \tag{14}$$

where:

$$R = (c^2-1)\sin^2\theta\, D_2^0\phi - \left(\frac{\kappa}{c} - 1\right)\sin 2\theta\, D_1^1\phi + \frac{\partial^2\phi}{\partial n^2} \tag{15}$$

with:

$$D_\kappa^\lambda = \left(-\sin\theta\frac{\partial}{\partial n} + \cos\theta\frac{\partial}{\partial s}\right)^{(\kappa)} \quad \left(\cos\theta\frac{\partial}{\partial n} + \sin\theta\frac{\partial}{\partial s}\right)^{(\lambda)} \tag{16}$$

Thus, relation (5) takes the form:

$$\frac{A_1}{c_1}\left[R_1 - B_1c_1\left(\frac{\partial^2\phi_1}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_1}{\partial n}\right)\right] = \frac{A_2}{c_2}\left[R_2 - B_2c_2\left(\frac{\partial^2\phi_2}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_2}{\partial n}\right)\right] \tag{17}$$

where R_i $(i = 1, 2)$ is obtained from R given from relation (15) by putting c_i for c and ϕ_i for ϕ.

Working similarly we obtain for the quantity:

$$\frac{\partial^2 v'}{\partial s^2} + \frac{1}{\rho}\frac{\partial u'}{\partial s} + u'\frac{\partial}{\partial s}\left(\frac{1}{\rho}\right)$$

which enters to both sides of relation (6) the following expression:

$$\frac{\partial^2 v'}{\partial s^2} + \frac{1}{\rho}\frac{\partial u'}{\partial s} + u'\frac{\partial}{\partial s}\left(\frac{1}{\rho}\right) = -A\left[\frac{\partial}{\partial n}(\nabla^2\phi) + (1+B)\frac{\partial}{\partial s}\left(\frac{\partial^2\phi}{\partial s\partial n}\right) - f\right] \tag{18}$$

where:

$$\begin{aligned} f = {} & \sin\theta\,(1+\cos^2\theta)\,(c-1)\,D_3^0\phi \\ & - \cos\theta[(1-c)\sin^2\theta + 2(\kappa-1)\cos^2\theta]D_2^1\phi \\ & + \sin\theta\left[\left(1-\frac{1}{c}\right)\cos^2\theta + 2(\kappa-1)\sin^2\theta\right]D_1^2\phi \\ & + \cos\theta(1+\sin^2\theta)\left(\frac{1}{c}-1\right)D_0^3\phi \\ & + \frac{1}{\rho}\left[(c-1)D_2^0\phi\cos 2\theta - \left(\frac{1}{c}-1\right)D_0^2\phi + 2(\kappa-1)D_1^1\phi\sin 2\theta\right] \end{aligned} \tag{19}$$

Thus relation (6) takes the form:

$$A_1\left[\frac{\partial}{\partial n}(\nabla^2\phi_1) + (1+B_1)\frac{\partial}{\partial s}\left(\frac{\partial^2\phi_1}{\partial s\partial n}\right) - f_1\right] = A_2\left[\frac{\partial}{\partial n}(\nabla^2\phi_2) + (1+B_2)\frac{\partial}{\partial s}\left(\frac{\partial^2\phi_2}{\partial s\partial n}\right) - f_2\right] \tag{20}$$

where f_i $(i = 1, 2)$ is obtained from f given from relation (19) by putting c_i for c and ϕ_i for ϕ.

Relations (17) and (20) constitute the displacement continuity conditions expressed in terms of the Airy stress functions ϕ_1 and ϕ_2 of the two media. For the particular case of isotropic materials ($c = 1$, $\kappa = 1$) these relations take the form:

$$A_1\left[\frac{\partial^2\phi_1}{\partial n^2} - B_1\left(\frac{\partial^2\phi_1}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_1}{\partial n}\right)\right] = A_2\left[\frac{\partial^2\phi_2}{\partial n^2} - B_2\left(\frac{\partial^2\phi_2}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_2}{\partial n}\right)\right] \tag{21}$$

and

$$A_1\left[\frac{\partial}{\partial n}(\nabla^2\phi_1) + (1+B_1)\frac{\partial}{\partial s}\left(\frac{\partial^2\phi_1}{\partial s\partial n}\right)\right] = A_2\left[\frac{\partial}{\partial n}(\nabla^2\phi_2) + (1+B_2)\frac{\partial}{\partial s}\left(\frac{\partial^2\phi_2}{\partial s\partial n}\right)\right] \tag{22}$$

Relations (21) and (22) were first obtained by Zienkiewicz and Gerstner [2].

Stress Continuity Conditions

The continuity of stresses along the interface AB (Fig. 1) is expressed by the following relations:

$$\sigma_n^{(1)} = \sigma_n^{(2)}, \quad \tau_{ns}^{(1)} = \tau_{ns}^{(2)} \tag{23}$$

Using transformation formulas:

$$\sigma_n = \frac{\partial^2\phi}{\partial x^2}\cos^2\theta + \frac{\partial^2\phi}{\partial y^2}\sin^2\theta + \frac{\partial^2\phi}{\partial x\partial y}\sin 2\theta = \frac{\partial^2\phi}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi}{\partial n}$$

$$\tau_{ns} = \frac{1}{2}\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right)\sin 2\theta - \frac{\partial^2\phi}{\partial x\partial y}\cos 2\theta = -\frac{\partial^2\phi}{\partial s\partial n} \tag{24}$$

we obtain from relations (23) the following relations:

$$\frac{\partial^2\phi_1}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_1}{\partial n} = \frac{\partial^2\phi_2}{\partial s^2} - \frac{1}{\rho}\frac{\partial\phi_2}{\partial n} \tag{25}$$

$$\frac{\partial^2\phi_1}{\partial s\partial n} = \frac{\partial^2\phi_2}{\partial s\partial n} \tag{26}$$

Relations (25) and (26) constitute the stress continuity conditions expressed in terms of the Airy stress functions ϕ_1 and ϕ_2 of the two media.

Examples

To obtain an idea of the usefulness of the previous derived displacement and stress continuity conditions we considered two particular problems: the two-material cantilever beam loaded by a concentrated force and couple at its free end, and the simply supported beam loaded by a uniform stress along its upper side.

(a) The Cantilever Beam. Consider a straight cantilever made of two orthotropic layers perfectly bonded along their interface (Fig. 2(*a*)). The free end of the cantilever is subjected to a transverse force P and a moment M. Let us assume that $E_x^{(1)} = E_1$, $E_x^{(2)} = 9E_1$, and that the heights of the two layers are equal to $(h/4)$ and $(3h/4)$, respectively. The thicknesses of both layers are taken equal to unity. Using the superscripts 1 and 2 for all quantities referred to the two layers we have for the boundary conditions the following relations:

$$\int_0^{h/4}\sigma_x^{(1)}(0,y)\,dy + \int_{h/4}^{h}\sigma_x^{(2)}(0,y)\,dy = 0$$

$$\int_0^{h/4}\sigma_x^{(1)}(0,y)\,y\,dy + \int_{h/4}^{h}\sigma_x^{(2)}(0,y)\,y\,dy = M$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\int_0^{h/4} \tau_{xy}(0,y)\ dy + \int_{h/4}^{h} \tau_{xy}(0,y)\ dy = -P$$

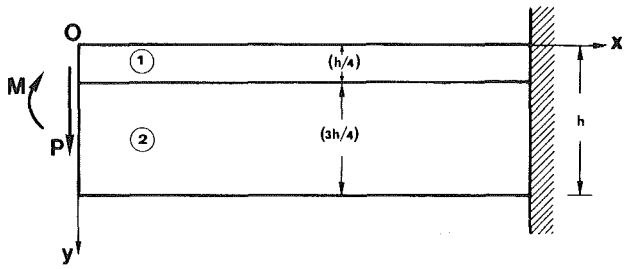
$$\tau_{xy}(x,0) = \sigma_y(x,0) = \tau_{xy}(x,h) = \sigma_y(x,h) = 0 \tag{27}$$

By assuming a stress function for the two layers of the form:

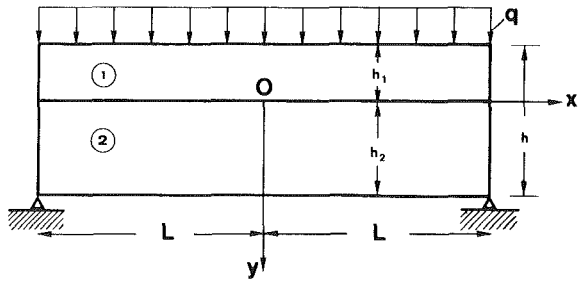
$$\phi_i = A_i\, xy + B_i\, y^2 + C_i\, xy^2 + D_i\, y^3 + E_i\, xy^3 \tag{28}$$

and using the boundary conditions (27) and the displacement and stress continuity conditions along the interface of the two layers expressed from relations (17), (20), and (25), (26) we obtain the following expressions for the stress components σ_x, σ_y, τ_{xy}:

$$\sigma_x^{(1)} = -\frac{204}{127h^2}(M-Px) + \frac{336}{127h^3}(M-Px)y$$



(a)



(b)

Fig. 2 (*a*) Geometry of a two-material composite cantilever; and (*b*) of a simply supported beam

$$\sigma_x^{(2)} = -\frac{918}{127h^2}(M-Px) + \frac{3024}{127h^3}(M-Px)y$$

$$\tau_{xy}^{(1)} = -\frac{204}{127}\frac{P}{h^2}y - \frac{168}{127}\frac{P}{h^3}y^2$$

$$\tau_{xy}^{(2)} = -\frac{324}{127}\frac{P}{h} - \frac{1836}{127}\frac{P}{h^2}y + \frac{1512}{127}\frac{P}{h^3}y^2 \tag{29}$$

Relations (29) coincide with those obtained by Lekhnitskii (reference [1], page 78).

(*b*) **The Simply Supported Beam.** Working similarly for the simply supported beam of Fig. 2(*b*) with elastic constants $E_x^{(i)}$, $\nu^{(i)}$, $G^{(i)}$ we obtain the following expressions for the stresses:

$$\sigma_x^{(1)} = -\lambda M(x)(2y + h_1 - a_1) + f_1(y)$$

$$\sigma_y^{(1)} = \lambda q\left[\frac{y^3}{3} + (h_1 - a_1)\frac{y^2}{2} - h_1 a_1 y + B_1\right]$$

$$\tau_{xy}^{(1)} = \lambda Q(x)\ (y + h_1)(y - a_1) \tag{30}$$

for the upper layer, and:

$$\sigma_x^{(2)} = -\mu M(x)\ (2y + a_2 - h_2) + f_2(y)$$

$$\sigma_y^{(2)} = \mu q\left[\frac{y^3}{3} + \frac{a_2 - h_2}{2}y^2 - h_2 a_2 y + B_2\right]$$

$$\tau_{xy}^{(2)} = \mu Q(x)\ (y - h_2)(y + a_2) \tag{31}$$

for the lower layer, with:

$$Q(x) = -qx$$

$$M(x) = \frac{q}{2}(L^2 - x^2) \tag{32}$$

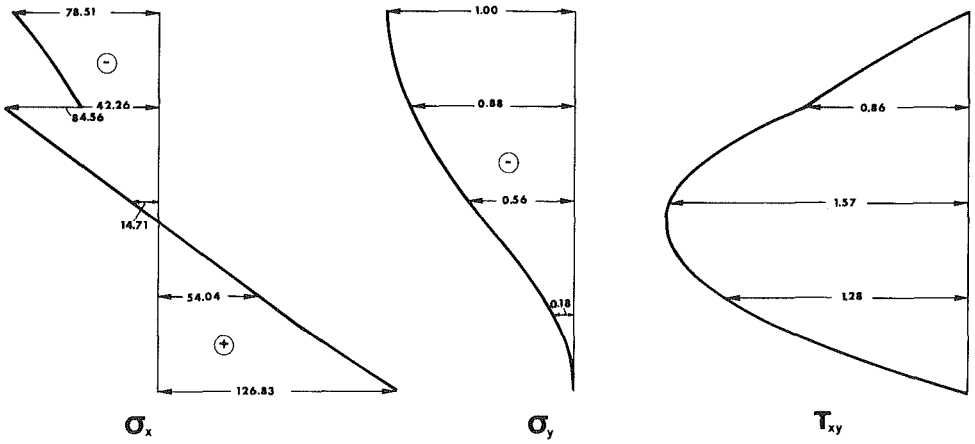
and:

$$f_1(y) = -\lambda q\left(\frac{E_x^{(1)}}{G^{(1)}} - 2\nu^{(1)}\right)\left(\frac{y^3}{3} + \frac{h_1 - a_1}{2}y^2\right) + \kappa_1 y + \kappa_2$$

$$f_2(y) = -\mu q\left(\frac{E_x^{(2)}}{G^{(2)}} - 2\nu^{(2)}\right)\left(\frac{y^3}{3} + \frac{a_2 - h_2}{2}y^2\right) + \kappa_1' y + \kappa_2'$$

$$a_1 = \frac{h_2 E_x^{(2)}}{h_1 E_x^{(1)} + h_2 E_x^{(2)}}h \qquad a_2 = \frac{h_1 E_x^{(1)}}{h_1 E_x^{(1)} + h_2 E_x^{(2)}}h \tag{33}$$

$$\lambda = -\frac{6E_x^{(1)}}{E_y} \qquad \mu = -\frac{6E_x^{(2)}}{E_y}$$



(a) (b) (c)

Fig. 3 Variation of the stress components σ_x, σ_y, τ_{xy} along the height of the beam of Fig. 2(*b*) for the cross section x = 1 m. All values are in Nm^{-2}.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$E_y = E_x^{(1)} h_1^3 + E_x^{(2)} h_2^3 + 3 \frac{E_x^{(1)} E_x^{(2)} h^2 h_1 h_2}{h_1 E_x^{(1)} + h_2 E_x^{(2)}}$$

$$B_1 = \frac{E_x^{(2)}}{E_x^{(1)}} \frac{h_2^2}{2} \left[\frac{h_1 E_x^{(1)} h}{h_1 E_x^{(1)} + h_2 E_x^{(2)}} + \frac{h_2}{3} \right]$$

$$B_2 = \frac{\lambda}{\mu} B_1$$

The coefficients κ_1, κ_2, κ_1', κ_2' are determined from the following relations:

$$\frac{\kappa_2'}{E_x^{(2)}} - \frac{\kappa_2}{E_x^{(1)}} = T_1, \quad \frac{\kappa_1'}{E_x^{(2)}} - \frac{\kappa_1}{E_x^{(1)}} = T_2$$

$$-\frac{\kappa_1}{2} h_1^2 + \kappa_2 h_1 + \kappa_1' \frac{h_2^2}{2} + \kappa_2' h_2 = S_1$$

$$\kappa_1 \frac{h_1^3}{3} - \kappa_2 \frac{h_1^2}{2} + \kappa_1' \frac{h_2^3}{3} + \kappa_2' \frac{h_2^2}{2} = S_2 \tag{34}$$

with:

$$S_1 = \mu\, q \left(\frac{E_x^{(2)}}{G^{(2)}} - 2\nu^{(2)} \right) \left(\frac{a_2 h_2^3}{6} - \frac{h_2^4}{12} \right) - \lambda q \left(\frac{E_x^{(1)}}{G^{(1)}} - 2\nu^{(1)} \right) \left(\frac{a_1 h_1^3}{6} - \frac{h_1^4}{12} \right)$$

$$S_2 = \mu\, q \left(\frac{E_x^{(2)}}{G^{(2)}} - 2\nu^{(2)} \right) \left(\frac{a_2 h_2^4}{8} - \frac{7}{120} h_2^5 \right) - \lambda q \left(\frac{E_x^{(1)}}{G^{(1)}} - 2\nu^{(1)} \right) \left(\frac{7 h_1^5}{120} - \frac{a_1}{8} h_1^4 \right) \tag{35}$$

$$T_1 = \frac{\nu_1^{(1)}}{E_x^{(1)}} \lambda B_1 - \nu_1^{(2)} \frac{\mu B_2}{E_x^{(2)}}$$

$$T_2 = -\left[\frac{1}{G^{(1)}} - \frac{\nu^{(1)}}{E_x^{(1)}} \right] \lambda q h_1 a_1 + \mu q h_2 a_2 \left(\frac{1}{G^{(2)}} - \frac{\nu^{(2)}}{E_x^{(2)}} \right)$$

As an application we considered the case:

$E_x^{(1)} = 0.105 \times 10^{11}\,\mathrm{Nm}^{-2}$ $\quad G^{(1)} = 0.006 \times 10^{11}\,\mathrm{Nm}^{-2}$, $\nu^{(1)} = 0.1$

$E_x^{(2)} = 0.210 \times 10^{11}\,\mathrm{Nm}^{-2}$ $\quad G^{(2)} = 0.01 \times 10^{11}\,\mathrm{Nm}^{-2}$, $\nu^{(2)} = 0.2$

$h_1 = 0.25$ m, $h_2 = 0.75$ m $\quad q = 1\ \mathrm{Nm}^{-1}$, $L = 6$m

Figure 3 presents the variation of the stresses σ_x, σ_y, τ_{xy} along the height of the beam for the cross section $x = 1.0$ m. A check of the equilibrium equations shows that they are satisfied with high accuracy.

Conclusions

In the present paper the stress and displacement continuity conditions along interfaces between two different bonded orthotropic media were formulated in terms of their Airy stress functions. These conditions were then used for the solution of two particular problems of two-material orthotropic beams. It was demonstrated that the obtained conditions greatly facilitate the solution of multilayered anisotropic problems by the Airy stress function method.

References

1 Timoshenko, S., and Goodier, J. N., *Theory of Elasticity,* McGraw-Hill, New York, 1951.

2 Lekhnitskii, G. C., *Anisotropic Plates,* translated from the Russian edition by Tsai, S. W., and Cheron, T., Cordon, Breach Sci. Publ., New York, 1968.

3 Zienkiewicz, O. C., and Gerstner, R. W., "A Stress Function Approach to Interface and Mixed Boundary-Condition Problems, Boundary Conditions and Finite-Difference Techniques," *International Journal of Mechanical Sciences,* Vol. 2, 1960, pp. 93–101.

4 Schile, R. D., "A Nonhomogeneous Beam in Plane Stress," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, Sept. 1962, pp. 582–583.

5 Schile, R. D., "Bending of Nonhomogeneous Bars," *International Journal of Mechanical Sciences,* Vol. 5, 1963, pp. 439–445.

6 Schile, R. D., "Laminate Elastostatics," *International Journal of Engineering Science,* Vol. 9, 1971, pp. 1–16.

7 Schile, R. D., "Analysis of Laminated Circular Plates," *Journal of Composite Materials,* Vol. 1, 1967, pp. 324–335.

8 Schile, R. D., "A Correspondence Principle for Laminated Plates," *International Journal of Engineering Science,* Vol. 12, 1974, pp. 569–584.

9 Reissner, E., and Stavsky, Y., "Bending and Stretching of Certain Types of Heterogeneous Aeolotropic Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, Sept. 1961, pp. 402–408.

10 Dong, S. B., Pister, K. S., and Taylor, R. L., "On the Theory of Laminated Anisotropic Shells and Plates," *Journal of the Aerospace Sciences,* Vol. 29, 1962, pp. 969–975.

11 Pagano, N. J., "Exact Solutions for Composite Laminates in Cylindrical Bending," *Journal of Composite Materials,* Vol. 3, 1969, pp. 398–411.

12 Gerstner, R. W., "Stresses in a Composite Cantilever," *Journal of Composite Materials,* Vol. 2, 1968, pp. 498–501.

13 Sierakowski, R. L., and Ebcioglu, I. K., "On Interlaminar Shear Stresses in Composites," *Journal of Composite Materials,* Vol. 4, 1970, pp. 144–149.

14 Whitney, J. M., Browning, C. E., and Mair, A., "Analysis of the Flexure Test for Laminated Composite Materials," in *Composite Materials: Testing and Design (Third Conference)*, ASTM STP 546, American Society for Testing and Materials, Philadelphia, Pa., 1974, pp. 30–45.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Radiation From Expanding Circular Dislocation Loops and Elastic Precursor Decay

X. Markenscoff
Associate Professor,
Department of Mechanical and Environmental Engineering,
University of California,
Santa Barbara, Calif. 93106

R. J. Clifton
Professor,
Division of Engineering,
Brown University,
Providence, R. I. 02912

The solution is given for a circular dislocation loop that instantaneously begins to expand by gliding in its plane at constant velocity. The material is assumed to be elastic and isotropic. The solution, expressed as an integral over the loop and over previous times, is evaluated explicitly near the wavefront. This wavefront solution is used to derive an expression for the decay in amplitude of an elastic plane wave propagating in the direction perpendicular to the plane of the loop due to the radiation of elastic waves from the expanding loops. By superimposing the effects of many loops set in motion as the plane wavefront passes, a relation is derived for the decay of the front of the plane wave. This precursor decay relation is found to be the same as obtained previously for infinitely long, straight dislocations based on consideration of either elastic waves radiated from discrete dislocations or energy dissipated in an elastic/viscoplastic continuum model of the material. Thus, dislocation line curvature can apparently be regarded as unimportant in the attenuation of the amplitude of the wavefront. Because detectors have finite rise times, measured precursor amplitudes are indicative of wave amplitudes at small but finite times Δt after the wavefront arrives. Evaluation of the effects of dislocation line curvature on wave amplitudes behind the wavefront requires the consideration of higher order terms in Δt which are neglected in the present analysis.

Introduction

Plate impact experiments are often used to study the dynamic plastic response of materials. One measure of this response is the decay of the leading elastic wave (elastic precursor) as it propagates through the thickness of the impacted plate. This so-called "precursor decay" is generally attributed to the inelastic response of the material corresponding to dislocations being set in motion by the stress state at the wavefront. Analyses of precursor decay in single crystals have been carried out both by modeling the crystals as elastic/viscoplastic continua and by considering explicitly the elastic waves emanating from dislocations that are set in motion by the stresses at the wavefront. Both of these analyses lead to essentially the same result: the precursor decay per unit distance of propagation is proportional to the product of average dislocation velocity and mobile dislocation density at the wavefront.

Comparison of measured and predicted precursor amplitudes indicates that the measured values are much less than those predicted using the initial total dislocation density and reasonable estimates for dislocation velocity. Various explanations of the discrepancy have been proposed; however, none has received general acceptance. A more complete discussion of the precursor decay anomaly and possible explanations is given in [1] and the references cited in [1]. The fundamental difficulty is that the dislocation density must be increased by 2–3 orders of magnitude during the rise time (less than 2×10^9 s) of the precursor. This is generally viewed as unlikely; however, it may be possible if dislocations can be regarded as loops expanding from a sufficiently small initial radius. To examine this possibility, the precursor decay relation based on expanding dislocation loops is required. Derivations of such a relation is the objective of this paper.

The analysis is similar in many respects to that given by Clifton and Markenscoff [1] for the case of infinitely long straight dislocations. For ease of analysis and interpretation, the stress wave incident on the dislocation loops is taken to be an elastic shear wave with the wavefront parallel to the plane of the loop. Based on the results in [1], the main features of precursor decay for this case are expected to be the same as for the case of the plate impact experiment in which there is an incident longitudinal wave with a wavefront inclined relative to the plane of the dislocation loops.

Several researchers have considered the motion of dislocation loops. Mura [2] appears to be the first to formulate the problem in terms of the dynamic Green's function for an impulsive force in a three-dimensional space. From a seismologist's point of view, Burridge and Knopoff [3]

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editoral Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July, 1981; final revision, March, 1982. Paper No. 82-WA/APM-19.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

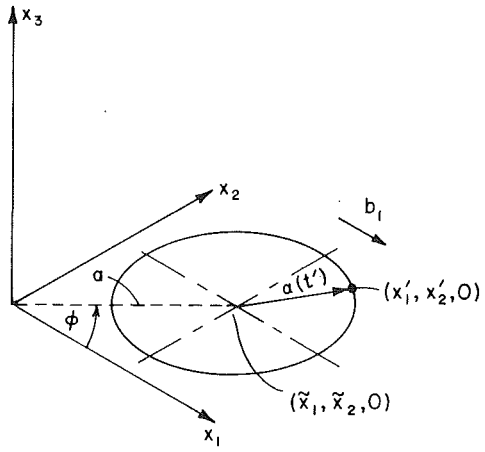


Fig. 1 Circular dislocation loop expanding in the plane $x_3 = 0$

formulated the problem in terms of surface integrals on the slip surface, Madariaga [4] treated an expanding circular fault numerically, while more recently Gu and Shuang [5] used spectral decomposition analysis to obtain the far field of an expanding circular dislocating loop, although without particular emphasis on the analysis of the singularities at the wavefront, which we obtain here. Günther [6] studied uniformly moving dislocations and considered a circular loop moving in a direction normal to its slip-plane. Mura [7] also studied the same motion but with the Burgers vector normal to the plane of the loop. The solution is given in terms of elliptic integrals and hypergeometric functions. The uniqueness of the solution of moving dislocation loops has been considered by Kossecka and Schöpf [8].

Here we consider a circular dislocation loop in an isotropic solid with the Burgers vector in the plane of the loop. The loop is at rest until time $t=0$ when it starts expanding with a radial velocity V, assumed constant. The solution for the velocity field is expressed in terms of the dynamic Green's function as a line integral along the loop and over the history of the motion. The asymptotic wavefront behavior is analyzed as it pertains to precursor decay analysis.

Radiation From an Expanding Dislocation Loop

Consider a circular dislocation loop (See Fig. 1) lying at rest in the plane $x_3 = 0$ with center at $(x_1, x_2, 0)$ and initial radius α_0. Without loss of generality the Burgers vector is taken along the x_1-direction with magnitude b_1. At time $t=0$ a shear pulse of stress amplitude σ_{13} propagating in the x_3-direction impinges on the dislocation, setting it in gliding motion with constant radial velocity V, i.e., the dislocation expands in its own plane as a circular loop.

The velocity field of a dislocation loop $\mathcal{L}(t)$, moving with velocity $V_i(\mathbf{x}', t')$ where $\mathbf{x}'$ denotes a point on the loop, is given by

$$\dot{U}_m(\mathbf{x},t) = \frac{\partial}{\partial x_l}\int_{-\infty}^{t} dt' \int_{\mathcal{L}(t')} C_{ijkl} G_{km}(\mathbf{x}-\mathbf{x}', t-t') V_r(\mathbf{x}', t') b_i \epsilon_{jrs} dl_s'(t') \tag{1}$$

where C_{ijkl} denotes the elastic coefficients, G_{km} is the Green's function for a unit impulse in a full space [9], $V_r(\mathbf{x},t')$ is the velocity vector at any point on the loop, b_i is the Burgers vector, and dl_s' is a line element along the loop; ϵ_{jrs} is the permutation symbol. The expression for the velocity field given in (1) differs from the expression given by Mura [2] (i.e., equation (20) in [2]) in that the differentiation is outside the integral, as required to obtain correct results. Kiusalaas and Mura [10] used the interpretation (1) in their application of (20) in [2].

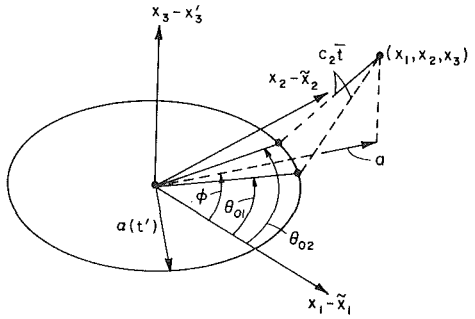


Fig. 2 Intersection of the domain of dependence of the solution at (x, t) with the dislocation loop at time t'

For isotropic materials, the Green's function is given by Love [9]:

$$G_{ij}(\mathbf{x}-\mathbf{x}', t-t') = \frac{1}{4\pi\rho}\left\{\frac{\bar{t}}{\bar{r}^2}\left(\frac{3\bar{r}_i\bar{r}_j}{\bar{r}^3} - \frac{\delta_{ij}}{\bar{r}}\right)H\left(\bar{t}-\frac{\bar{r}}{c_1}\right)H\left(\frac{\bar{r}}{c_2}-\bar{t}\right)\right.$$

$$\left. + \frac{\bar{r}_i\bar{r}_j}{\bar{r}^3}\left[\frac{1}{c_1^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_1}\right) - \frac{1}{c_2^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_2}\right)\right] + \frac{\delta_{ij}}{\bar{r}c_2^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_2}\right)\right\} \tag{2}$$

where ρ is the density, $\bar{t}=t-t'$, $\bar{r}_i = x_i - x_i'$, $r^2 = x_1^2 + x_2^2 + x_3^2$, $\bar{r}^2 = (x_1-x_1')^2 + (x_2-x_2') + (x_3-x_3')^2$; c_1 and c_2 are the longitudinal and shear wave speeds, $H(\ )$ denotes the Heaviside step function, $\delta(\ )$ denotes the delta function, and δ_{ij} is the Kronecker delta. For the slip geometry being considered, and for isotropic materials, the only nonzero terms in (1) are obtained from the terms containing

$$C_{1313}G_{11} \quad \text{and} \quad C_{1331}G_{31}. \tag{3}$$

The coordinates of a point on a circular loop with initial radius α_0 that begins to expand on its own plane with constant radial velocity V at time $t' = 0$ are

$$\begin{aligned} x_1' &= \tilde{x}_1 + \alpha(t')\cos\theta \\ x_2' &= \tilde{x}_2 + \alpha(t')\sin\theta \\ x_3' &= 0 \end{aligned} \tag{4}$$

where $\alpha(t') = \alpha_0 + Vt'H(t')$.

From (1)–(4) the velocity component $\dot{U}_1$ can be written as

$$\dot{U}_1(x_1,x_2,x_3,t) = \frac{\partial}{\partial x_3}\int_0^t dt' \int_0^{2\pi} b_1 \frac{C_{1313}}{4\pi\rho}\left[\left(-\frac{\bar{r}_1^2}{\bar{r}^3} + \frac{1}{\bar{r}}\right)\frac{1}{c_2^2}\right.$$

$$\left.\delta\left(\bar{t}-\frac{\bar{r}}{c_2}\right) + \frac{\bar{r}_1^2}{\bar{r}^3}\frac{1}{c_1^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_1}\right) + \frac{3\bar{t}\bar{r}_1\bar{r}_3}{\bar{r}^5}H\left(\bar{t}-\frac{\bar{r}}{c_1}\right)H\left(\frac{\bar{r}}{c_2}-\bar{t}\right)\right]$$

$$V\alpha(t')d\theta + \frac{\partial}{\partial x_1}\int_0^t dt' \int_0^{2\pi} b_1 \frac{C_{1331}}{4\pi\rho}\left[-\frac{\bar{r}_1\bar{r}_3}{c_2^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_2}\right)\right.$$

$$\left. + \frac{\bar{r}_1\bar{r}_3}{\bar{r}^3}\frac{1}{c_1^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_1}\right) + \frac{3t\bar{r}_1\bar{r}_3}{\bar{r}^5}H\left(\bar{t}-\frac{\bar{r}}{c_1}\right)H\left(\frac{\bar{r}}{c_2}-\bar{t}\right)\right]$$

$$V\alpha(t')d\theta. \tag{5}$$

First we consider the integral

$$\frac{\partial}{\partial x_3}\int_0^t dt' \int_0^{2\pi} b_1 \frac{C_{1313}}{4\pi\rho}\left(-\frac{\bar{r}_1^2}{\bar{r}^3} + \frac{1}{\bar{r}}\right)\frac{1}{c_2^2}\delta\left(\bar{t}-\frac{\bar{r}}{c_2}\right)V\alpha(t')d\theta \tag{6}$$

and carry out the integration with respect to θ:

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\int_0^{2\pi}\left(-\frac{\tilde{r}_1^{\,2}}{\tilde{r}^3}+\frac{1}{\tilde{r}}\right)\delta\left(\tilde{t}-\frac{\tilde{r}}{c_2}\right)d\theta=\sum_{\theta_0}\left[(1-\tilde{r}_1^2/\tilde{r}^2)/\right.$$

$$\left.(|df/d\theta|\tilde{r})_{\theta=\theta_0}\right. \tag{7}$$

where

$$f=f(\theta;t')\equiv\tilde{t}-\frac{\tilde{r}}{c_2} \tag{8}$$

and θ_0 are the zeros of $f(\theta;t')$. These zeros correspond to the points where the domain of dependence of the solution at (**x**, t) intersects the loop at time t' (See Fig. 2). From (4) and (8) the values θ_0 satisfy

$$\cos\phi\cos\theta_0+\sin\phi\sin\theta_0=\pm C \tag{9}$$

in which

$$\phi=\tan^{-1}\frac{x_2-\tilde{x}_2}{x_1-\tilde{x}_1} \tag{10a}$$

and

$$C=\frac{1}{2\alpha(t')a}[c_2^2\tilde{t}^2-\alpha(t')^2-a^2-x_3^2] \tag{10b}$$

where

$$a=[(x_1-\tilde{x}_1)^2+(x_2-\tilde{x}_2)^2]^{1/2}. \tag{10c}$$

The + sign in (9) applies for $0\le\phi\le\pi/2$ and the − sign for $\pi/2\le\phi\le\pi$. The two values of θ_0 that satisfy (9) are

$$\theta_{01,02}=\phi\pm\cos^{-1}C. \tag{11}$$

The angle ϕ corresponds to the angle that the plane normal to $x_3=0$ and containing (x_1,x_2,x_3) and the center of the loop makes with the x_1 axis. The angles $\theta_{01,02}$ locate the intersection of the sphere with center (x_1,x_2x_3) and radius $c_2\tilde{t}$ with the loop at time t' (see Fig. 2).

From (4) and (8) we obtain

$$\left.\frac{df}{d\theta}\right|_{\theta=\theta_0}=\frac{\alpha(t')}{c_2\tilde{r}}\{(x_1-\tilde{x}_1)\sin\theta_0-(x_2-\tilde{x}_2)\cos\theta_0\} \tag{12}$$

which, by means of (10a) and (10b) can be rewritten as

$$\left.\frac{df}{d\theta}\right|_{\theta=\theta_0}=\frac{\alpha(t')a}{c_2^2\tilde{t}}\sin(\theta_0-\phi). \tag{12$'$}$$

Finally, using (11) we obtain

$$\left.\frac{df}{d\theta}\right|_{\theta=\theta_0}=\pm\frac{\alpha(t')a}{c_2^2\tilde{t}}\sqrt{1-C^2}. \tag{12$''$}$$

For the angles θ_0 in (11) to be real, the values of C in (10b) must satisfy

$$-1\le C\le 1. \tag{13}$$

This defines the upper limit of the integration in t'. From (10b) and (4) it follows that the condition (13) is

$$-c_2^2(t-t')^2+(\alpha_0+Vt')^2+a^2+x_3^2-2(\alpha_0+Vt')a\le 0. \tag{14}$$

The range of t' for which (15) holds is $t'\le t_{01}$ and $t'\ge t_{02}$, where $t_{01,02}$ are the zeros of (14); i.e.,

$$(V^2-c_2^2)t_0^2+2V\left(\alpha_0-a+\frac{c_2^2t}{V}\right)t_0+(\alpha_0-a)^2-c_2^2t^2+x_3^2=0 \tag{15}$$

For subsonic motion (15) has two positive roots, say t_{01} and t_{02} with $t_{01}<t_{02}$. The larger root, t_{02}, is greater than t; thus, the range of integration in t' is $[0,\ t_{01}]$. Solving (16), we obtain

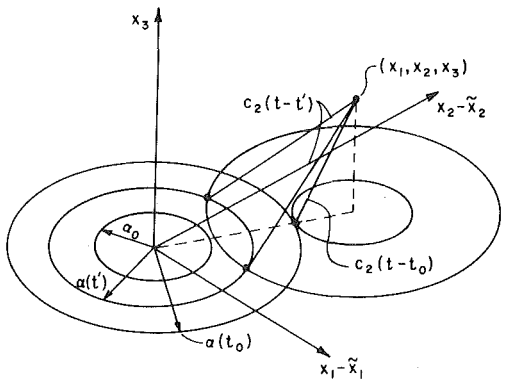


Fig. 3 Upper limit $t'=t_0$ for contribution to the solution at (x, t)

$$t_{01}=\frac{V\left(\alpha_0-a+\frac{c_2^2t}{V}\right)-c_2\left[\left((a-\alpha_0)-Vt\right)^2+(1-\frac{V^2}{c_2^2})x_3^2\right]^{1/2}}{c_2^2-V^2}. \tag{16}$$

The geometrical interpretation of (16) is shown in Fig. 3. The sphere with center at (x_1,x_2,x_3) and radius $c_2(t-t')$ intersects the plane of the loop along a circle. The intersection of this circle with the loop defines the dislocation segment which at t' emanates wavelets that contribute to the solution at the field point (x_1,x_2,x_3) at time t. From Fig. 3, the latest time $t'=t_{01}$ for which the loop contributes to the solution at (x_1,x_2,x_3) at time t satisfies

$$(\alpha_0+Vt_{01})+\sqrt{c_2^2(t-t_{01})^2-x_3^2}=a, \tag{17}$$

which may be solved for t_{01} to obtain (16).

Finally, substituting (12)″ in (7) we obtain (6) in the form

$$\frac{C_{1313}}{2\pi\rho}b_1\frac{V}{c_2^2}\frac{\partial}{\partial x_3}\int_0^{t_{01}}\left[-\frac{(x_1-\tilde{x}_1-\alpha(t')\cos\theta_0)^2}{c_2^3\tilde{t}^3}+\frac{1}{c_2\tilde{t}}\right]\frac{c_2^2\tilde{t}}{a(1-C^2)^{1/2}}dt' \tag{18}$$

where t_{01} is defined by (16). The remaining terms in (5) can be integrated with respect to θ similarly. However, this is unnecessary for the objective of this paper since the main contribution comes from the term (6) or, equivalently, (18).

Evaluation Near Wavefront

For the elastic precursor analysis we are interested in evaluating the particle velocity $\dot{U}_1$ at $(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ due to a dislocation loop that began expanding at the point $(x_1,\ x_2,\ x_3')$ on the plane $x_3=x_3'$ at time $t=x_3'/c_2$. The segments of the loop that contribute to the solution at $(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ must, at time x_3'/c_2, be within the distance $c_2(\hat{x}_3/c_2+\Delta t-x_3'/c_2)$ of the point $(0,0,\hat{x}_3)$. Such segments lie inside a paraboloid of revolution about the x_3 axis as shown in Fig. 4. The equation of this paraboloid is

$$2(\hat{x}_3-x_3)c_2\Delta t=(x_1^2+x_2^2)-c_2^2(\Delta t)^2 \tag{19}$$

which is analogous to equation (7) of [1]. The distance $|a-\alpha_0|=(x_1^2+x_2^2)^{1/2}$ from the x_3-axis to dislocation segments which contribute to the solution at $(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ is from (19), $0(\sqrt{\Delta t})$. From (16) and (19), the latest time $t'=t_{01}$, at which a loop expanding on the plane $x_3=0$ contributes to the solution at $(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ is

$$t_{01}=\Delta t-\frac{\hat{x}_3}{2c_2}\omega^2+\text{higher order terms in }\Delta t\text{ and }\omega \tag{20}$$

where (see Fig. 4)

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

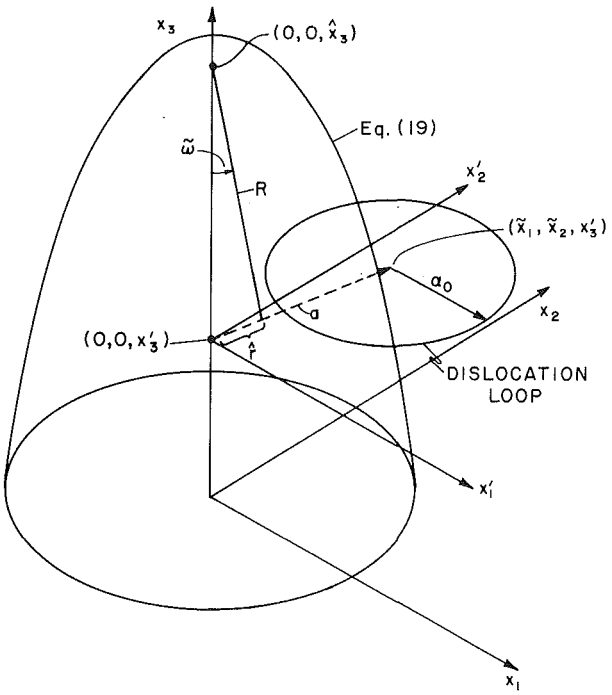


Fig. 4 Dislocation loops contributing to the solution at $(0,0,\hat{x}_3, \hat{x}_3/c_2 + \Delta t)$

$$\omega \equiv (a-\alpha_0)/\hat{x}_3 \simeq \tan^{-1}[(a-\alpha_0)/\hat{x}_3]$$

and ω is $0(\sqrt{\Delta t})$ from (19). For elastic precursor analysis the expression (20) is used for the upper limit t_{01} in the evaluation of (18).

In evaluating the contribution to the value of $\dot{U}_1$ at $(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ we note, from (10*b*) and (17), that t_{01} is a root of $(1-C^2)$; hence, the order of differentiation and integration shown in (18) should be maintained. From (10*b*) the quantity $(1-C^2)$ can be written as

$$1-C^2 = \frac{1}{4a^2\alpha^2}\left[(\alpha+a)^2 - c_2^2\bar{t}^2 + \hat{x}_3^2\right]\left[-(\alpha-a^2) + c_2^2\bar{t}^2 - \hat{x}_3^2\right] \tag{21}$$

where $\alpha = \alpha(t')$. For $t = \hat{x}_3/c_2 + \Delta t$ and t' in the interval $(0, t_{01})$ with t_{01} given by (20) the term $(c_2^2\bar{t}^2 - \hat{x}_3^2)$ is negligible in comparison with $(\alpha + a)^2$, if it is assumed that the loop radius α is much greater than $c_2\Delta t$; furthermore, $\alpha \to a$ as $\Delta t \to 0$. Thus, for evaluation of (18) at $(0, 0, \hat{x}_3, \hat{x}_3/c_2 + \Delta t)$ as $\Delta t \to 0$ we can approximate (21) by

$$1-C^2 \simeq \frac{1}{a^2}\left[c_2^2\bar{t}^2 - \hat{x}_3^2 - (\alpha-a)^2\right]. \tag{21}'$$

Considering the second term in (18) and using (21)′ we obtain

$$\dot{U}_1^I \simeq \frac{C_{1313}\, b_1 V}{2\pi\rho c_2}\frac{\partial}{\partial \hat{x}_3}\int_0^{t_{01}} \frac{dt'}{\sqrt{X(t')}} \tag{22}$$

where $\dot{U}_1^I$ denotes the contribution from the second term in (6), to the value of $\dot{U}_1(0,0,\hat{x}_3,\hat{x}_3/c_2+\Delta t)$ as $\Delta t \to 0$ and

$$X(t') \equiv Dt'^2 + Bt' + A \tag{23}$$

in which

$$D = c_2^2 - V^2$$
$$B = -2c_2^2 t - 2V(\alpha_0 - a)$$
$$C = c_2^2 t^2 - (\alpha_0 - a)^2 - \hat{x}_3^2.$$

Integration of (22) gives

$$\dot{U}_1^I = \lim_{\Delta t \to 0} \frac{C_{1313} b_1 V}{2\pi\rho c_2}\left\{\frac{1}{D^{1/2}}\log\left[\frac{\sqrt{X(t_{01})} + D^{1/2}t_{01} + B/(2D^{1/2})}{\sqrt{X(0)} + B/(2D^{1/2})}\right]\right\} \tag{24}$$

The term $X(t_{01})$ in (24) is zero because t_{01} is defined to be a root of $X(t')$. The quantity $X(0)$ is

$$X(0) = c_2^2 t^2 - (\alpha_0 - a)^2 - \hat{x}_3^2 \tag{25}$$

which, for small Δt becomes

$$X(0) \simeq \hat{x}_3^2\left(\frac{2c_2\Delta t}{\hat{x}_3} - \omega^2\right) \tag{25}'$$

Differentiation of (24) and use of $\partial/\partial\hat{x}_3[X(0)] = -2\hat{x}_3$ gives

$$\dot{U}_1^I \simeq \frac{C_{1313} b_1 V}{2\pi\rho c_2 D^{1/2}}\left\{\frac{D^{1/2}(1-\omega^2/2)}{c_2(L - D^{1/2}t_{01})} - \frac{\hat{x}_3}{(L-\sqrt{X(0)})\sqrt{X(0)}}\right\} \tag{26}$$

where

$$L = \frac{c_2^2 t + V(\alpha_0 - a)}{D^{1/2}}.$$

As $\Delta t \to 0$ the quantities ω, t_{01}, and $X(0)$ in (26) vanish and $L \to c_2\hat{x}_3/D^{1/2}$. Thus, the first term in { } in (26) remains finite as $\Delta t \to 0$, but the second term becomes infinite because $X(0) \to 0$. Consequently, the second term is dominant for sufficiently small Δt and we obtain, after using $C_{1313} = \rho c_2^2$,

$$\dot{U}_1^I \simeq \frac{-b_1 V}{2\pi\hat{x}_3[2c_2\Delta t/\hat{x}_3 - \omega^2]^{1/2}} \tag{27}$$

as $\Delta t \to 0$.

The remaining contribution from (18) vanishes as $\Delta t \to 0$ since the numerator contains the term

$$(\tilde{x}_1 + \alpha(t')\cos\theta_0)^2 = [-a\cos\phi + \alpha(t')(C\cos\phi - (1-C^2)^{1/2}\sin\phi)]^2 \tag{28}$$

which, at $t' = t_{01}$, cancels the singularity from the term $(1-C^2)^{-1/2}$. Moreover, it can be shown that the remaining integrals in (5) give contributions that vanish as $\Delta t \to 0$. Thus, the only term that contributes to precursor decay is (27).

Superpositon to Obtain Precursor Decay

Consider a plane shear wave

$$\dot{U}_1(\mathbf{x},t) = V_0 H(t - x_3/c_2), \tag{29}$$

with V_0 constant, which causes dislocation loops to expand (or contract) as the wavefront passes. As these loops expand elastic waves are radiated which tend to reduce the value of $\dot{U}_1$ near the wavefront. This reduction can be calculated by using the near wavefront solution (27) and superimposing the contributions from all dislocations that have segments within the domain of dependence of the solutions at $(0,0,\hat{x}_3)$ at time $t = \hat{x}_3/c_2 + \Delta t$. From (27) the contribution to $\dot{U}_1(0,0,\hat{x}_3,\hat{x}_3/c_2 + t)$ due to a dislocation loop in the plane $x_3 = x_3'$ that begins to expand with velocity V at time x_3'/c_2 is, from (27),

$$\dot{U}_1 = -\frac{b_1 V}{2\pi(\hat{x}_3 - x_3')[2c_2\Delta t/(\hat{x}_3 - x_3') - \tilde{\omega}^2]^{1/2}} \tag{30}$$

in which

$$\tilde{\omega} = \frac{\alpha_0 - a}{(\hat{x}_3 - x_3')}$$

where α_0 and a are, respectively, the radius of the dislocation loop in the plane $x_3 = x_3'$ and the distance from the center of this loop to the x_3-axis (see Fig. 4).

In order to sum the contributions of the form (30) from the

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

various loops that contribute to the solution at $(0,0,\hat{x}_3, \hat{x}_3/c_2 + \Delta t)$, it is helpful to note that a contribution of the form (30) arises from every dislocation loop that intersects a backward-drawn characteristic ray through this point (see Fig. 2). For Δt sufficiently small the maximum radius $(2\hat{x}_3 c_2 \Delta t)^{1/2}$ (See (19)) of the paraboloid of revolution in Fig. 4 is less than any finite value of dislocation loop radius. Consequently, as Δt approaches zero the number of dislocation loops that intersect the rays through $(0,0,\hat{x}_3,\hat{x}_3/c_2 + \Delta t)$ becomes equal to the number of dislocation loops with centers outside the paraboloid of revolution in Fig. 4 and with segments passing through the enclosed region. We begin by deriving an expression for the number of loops that pass through an annular volume element of this region. We assume that all dislocation loops have radius α_0.

If N_L is used to denote the number of dislocation loops per unit volume, then the number of centers that are located in an annulus of radius r, height dx_3', and width dr is

$$dM_r = 2\pi N_L r dr dx_3'. \tag{31}$$

In order for a dislocation loop of radius α_0 to have a segment within a circle of radius $\hat{r} + d\hat{r}$ but not within a circle of radius $\hat{r}$ (centered on the x_3-axis) the center of the dislocation loop must be in an annulus of width $d\hat{r}$ of a circle centered on the x_3-axis and having radius $\alpha_0 + \hat{r}$ or $\alpha_0 = \hat{r}$, for dislocation loop radii $\alpha_0 > \hat{r}$. From (31) the number of segments that pass through the annulus of width $d\hat{r}$ is

$$dN_{\hat{r}} = 2\pi N_L (\alpha_0 + \hat{r}) d\hat{r} dx_3' + 2\pi N_L (\alpha_0 - \hat{r}) d\hat{r} dx_3'$$

or

$$dN_{\hat{r}} = 4\pi N_L \alpha_0 d\hat{r} dx_3'. \tag{32}$$

Using $\hat{r} = R \sin \tilde{\omega}$ (see Fig. 4) and differentiating, we can rewrite (32) as

$$dN_{\hat{r}} = 4\pi N_L \alpha_0 R \cos \tilde{\omega} d\tilde{\omega} dx_3' = 2N(\hat{x}_3 - x_3') d\tilde{\omega} dx_3' \tag{32}'$$

where $N = 2\pi\alpha_0 N_L$ is the dislocation line length per unit volume. Equation (32)′ is the basic expression for the number of loops that pass through an annular volume element of the paraboloid of revolution that represents the domain of dependence of the solution at $(0,0,\hat{x}_3,\hat{x}_3/c_2 + \Delta t)$.

Superposition of the contributions from all dislocations within the domain of dependence of the solution at $(0,0,\hat{x}_3,\hat{x}_3/c_2 + \Delta t)$ gives, from (30) and (32)′

$$\dot{U}_1 = \frac{-b_1 V}{\pi} \int_0^{\hat{x}_3 - R_{III}} N dx_3' \int_0^{\tilde{\omega}_{max}} \frac{d\tilde{\omega}}{[\tilde{\omega}_{max} - \tilde{\omega}^2]^{1/2}} \tag{33}$$

where $\tilde{\omega}_{max} = 2c_2 \Delta t / (\hat{x}_3 - x_3')$ and R_{III} is an arbitrary small distance that is sufficiently large relative to $c_2 \Delta t$ to insure the validity of the near wavefront approximation (27) used in deriving (33). Contributions from the region $\hat{x}_3 - x_3' < R_{III}$ are taken to be negligible based on the behavior established in [1] for infinitely long, straight dislocations. Integration of the rightmost integral in (33) gives the value $\pi/2$. Then, (33) becomes

$$\dot{U}_1 = \frac{-b_1 V}{2} \int_0^{\hat{x}_3 - R_{III}} N dx_3'. \tag{34}$$

Differentiation of (34) with respect to $\hat{x}_3$ gives the precursor decay relation

$$\frac{D\dot{U}_1}{D\hat{x}_3} = \frac{-b_1 NV}{2}, \tag{35}$$

which is the same as the one obtained in [1] for the case of infinitely long, straight screw dislocations on planes perpendicular to the direction of propagation of a plane shear wave (cf. (35) with (17) of [1] for $\psi = 0$).

Concluding Remarks

The precursor decay relation for dislocation loops of uniform initial radius has been shown to be independent of the initial loop radius, provided that the loop radius is large compared with $c_2 \Delta t$. This result suggests that N, the dislocation line length per unit volume, is the only geometric property of dislocation loops that is important in precursor decay when precursor amplitude is interpreted as amplitude at the wavefront (i.e., for $\Delta t = 0$). Thus, provided that the time resolution of the detectors allows the wavefront amplitude to be determined accurately, it appears that discrepancies between observed precursor decay and that predicted on the basis of relations of the form (35) is not due to an inadequacy of the latter relations for the complex loop configurations that occur in real crystals. The finite response time of detectors may however be a critical factor since computed wave profile show a sharp decrease in amplitude immediately behind the wavefront. Further analysis of the near wavefront solution for small but finite Δt would be helpful in determining the effect of dislocation loop curvature on the amplitude behind the wavefront.

Acknowledgment

This work was supported by the National Science Foundation through individual grants to each author and by the NSF Materials Research Laboratory at Brown University.

References

1 Clifton, R. J., and Markenscoff, X., "Elastic Precursor Decay and Radiation From Nonuniformly Moving Dislocations," *Journal of the Mechanics and Physics of Solids*, Vol. 29, 1981, pp. 227–251.

2 Mura, T., "Continuous Distribution of Moving Dislocations," *Philosophical Magazine*, Vol. 8, 1963, pp. 843–856.

3 Burridge, R., and Knopoff, L., "Body Force Equivalent for Seismic Dislocations," *Bulletin of the Seismological Society of America*, Vol. 54, 1964, pp. 1875–1888.

4 Madariaga, R., "Dynamics of an Expanding Circular Fault," *Bulletin of the Seismological Society of America*, Vol. 66, 1976, pp. 639–666.

5 Gu, J., and Shuang, L., "Far Field Theory Radiated From a Circular Dislocation Expanding with Uniform Varying Velocity," Part I, Part II (unpublished manuscript).

6 Gunther, H., "Uberschallbewegung von Eigenspannungsquellen in der Kontinuumstheorie," *Annalen der Physik*, Vol. 21, 1968, pp. 93–105.

7 Mura, T., "Methods of Continuously Distributed Dislocations," in *Mathematical Theory of Dislocations*, Mura, T., ed., ASME, 1969, pp. 25–48.

8 Kossecka, E., and Schopf, H. G., "On the Uniqueness Problem of a Dislocation Field," *Archiwum Mechaniki Stosowanej*, Vol. 24, 1972, pp. 27–34.

9 Love, E. H., *The Mathematical Theory of Elasticity*, Dover, New York, 1944, p. 305.

10 Kiusalaas, J., and Mura, T., "On the Elastic Field Around an Edge Dislocation With Application to Dislocation Vibration," *Philosophical Magazine*, Vol. 9, 1964, pp. 1–7.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Wavefront Analysis in the Nonseparable Elastodynamic Quarter-Plane Problems, I

Part 1: The General Method

J. Miklowitz
Professor,
Division of Engineering
and Applied Science,
California Institute of Technology,
Pasadena, Calif. 91125
Fellow ASME

This work concerns the transient response in elastodynamic quarter-plane problems. In particular, the work focuses on the nonseparable problems, developing and using a new method to solve basic quarter-plane cases involving nonmixed edge conditions. Of initial interest is the classical case in which a uniform step pressure is applied to one edge, along with zero shear stress, while the other edge is traction free. The new method of solution is related to earlier work on nonseparable wavequide (semi-infinite layer) problems in which long time, low-frequency response was the objective. The basic idea in the earlier work was the exploitation of a solution boundedness condition in the form of an infinite set of integral equations. These equations were generated by eliminating the residues at an infinite set of poles stemming from the zeros of the Rayleigh-Lamb frequency equation. These poles were associated with exponentially unbounded contributions. The solution of the integral equations determined the unknown edge conditions in the problem, hence giving the long time solution. In the new method for the quarter-plane problems the analog of the Rayleigh-Lamb frequency equations is the Rayleigh function that yields just three zeros, i.e., say the squares of c_{R1}, c_{R2}, and c_{R3} speeds for the Rayleigh surface waves. Of these six roots (and corresponding poles) only one is physical, i.e., c_{R1}. The other five are eliminated since they correspond to: (1) exponentially unbounded waves, or (2) waves with speeds $c_{R2} > c_d$, $c_{R3} > c_d$, all inadmissible. They analogously (to the layer case) give four integral equations for the edge unknowns, hence the formal solution for the problems. To date the method has been successful in obtaining all the wavefront expansions for the uniform step pressure case mentioned earlier. The dilatational and equivoluminal wavefronts were obtained for the interior and both edges and the Rayleigh wavefronts for both edges. This asymptotic solution has been verified and hence establishes the credibility of the new method.

Introduction

The general problem of wave propagation in a two-dimensional elastic wedge with nonmixed edge conditions (i.e., both stresses or both displacements specified) is one of considerable difficulty but of long standing interest and importance in mechanics and seismology. The basic difficulty in getting solutions is due to the nonseparability of such problems. In an excellent survey published in 1969 Knopoff [1] reviewed the various analytical methods that have been tried on these problems. The survey ends on the sad note that none of the techniques tried (images, integral transforms, Weiner-Hopf method, and others) were successful. In 1969 Alterman and Rotenberg [2] did carry out a successful numerical analysis of the quarter-plane with free edges, excited by a buried source.

Very recently there have been some good gains made with methods based on self-similar solutions which are applicable to problems not involving a characteristic length. Achenbach and Khetan [3], for example, have obtained a solution based on a system of integral equations, for the problem of the wedge subjected to edge pressures. Achenbach [4] has

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted uitil two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1981; final revision, April, 1982. Paper No. 82-WA/APM-12.

Copies will be available unitl July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

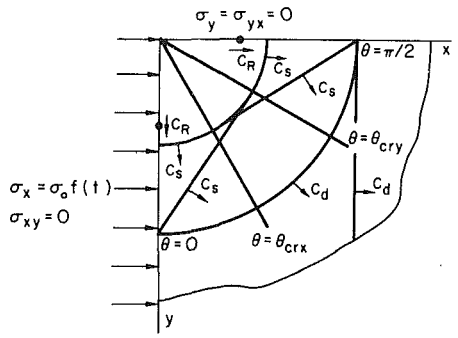


Fig. 1 The elastodynamic quarter-plane problem, its wavefronts, and critical shear regions, $0 \le \theta \le \theta_{cry}$ and $\theta_{crx} \le \theta \le \pi/2$

reported numerical results for the radial particle velocity along the wedge edges for the case of antisymmetical traveling line loads on these edges. Wojcik [5], also using self-similar solutions, has been able to reduce this type of wedge problem (no characteristic length) to a single unknown governed by a Fredholm integral equation of the second kind in one dimension. To date numerical evaluation of this equation, needed to obtain wave information in the problem, has not been carried out.

The present paper is concerned with the quarter-plane problem, an important special case of the preceding wedge problems, where the vertex angle is $\pi/2$. Lapwood [6], using a triple transform method, studied such a problem in which one edge of the quarter plane was subjected to a normal line load a distance, say a, from the vertex under zero shear stress. He found an approximation for the Rayleigh surface pulse that had rounded the corner. Indeed, changes in the Rayleigh pulse, as it negotiates a general angular discontinuity (i.e., transmission and reflection), have been of great concern in the seismology and mechanics literature. The problem was studied early by Viktorov, de Bremaecher, Knopoff and Gangi, and others by analytical (through iteration methods) and experimental techniques. Reference [1] cites these references. Still others may be found in the survey by Miklowitz [7].

Using the basic theme of recent work in nonseparable elastic waveguide problems (i.e., those with nonmixed edge conditions) by Miklowitz and coworkers [8–10], exploiting a boundedness condition on the solution, the author has found a similar boundedness condition on solutions, and a related condition that eliminates inadmissible events in such solutions, for the nonseparable quarter-plane problems. The waveguide and quarter-plane problems are, of course, quite different, the former having a characteristic length and the latter not. This leads to certain distinct differences in the boundedness condition and other parts of the method for these two different types of problems, one (the waveguide) being an eigenvalue problem and the other (the quarter-plane) not.

To outline the present general technique for determining the wavefronts in nonseparable quarter-plane problems we draw on the pressure shock problem depicted in Fig. 1. In-plane coordinates are x and y with u and v corresponding displacement components, respectively. Note the nonmixed stress-type edge conditions. The normal stress $\sigma_x = \sigma_0 f(t)$ is suddenly applied to $t = 0$ to edge $x = 0$ under zero shear stress. Aside from being a vehicle for demonstrating the solution technique, the problem is a long standing basic one in elastodynamics. The present work determines all the wavefronts for the waves in the problem. Figure 1 depicts these wave events, interior and on the edges (surfaces). Note that there are two critical shear regions, i.e., $\theta = \theta_{crx}$ and $\theta = \theta_{cry}$. They are associated with the x and y-axes, respectively.

These problems, like the waveguide ones, with their nonmixed edge conditions cannot be solved by direct separation techniques. The technique here employs the plane-strain displacement equations of motion, transforming out time t with the one-sided Laplace transform (parameter p) in the usual way. There remains the need for a procedure to reduce the coordinate x, leaving coordinate y and ordinary differential equations in this variable. However, no transforms on x exist that "ask for" the nonmixed edge displacements or stresses, and *therein lies the problem.* The present method proceeds by employing a one-sided Laplace transform on x (parameter s) which gives rise to *time-transformed edge unknowns* (on $x = 0$) for the displacements and their gradients, e.g., $\bar{u}(0,y,p)$, $\bar{u}_x(0,y,p)$, etc. Then employing the traction-free conditions on the edge $y = 0$, one generates a generalized form of the Rayleigh function $R(s,p) = (k_s^2 - 2s^2)^2 + 4s^2\alpha\beta$ where $\alpha = i(s^2 - k_d^2)^{1/2}$ and $\beta = i(s^2 - k_s^2)^{1/2}$ with p and s complex, and $k_d = p/c_d$, $k_s = p/c_s$, c_d and c_s being the dilatational and equivoluminal body wave speeds. $R(s,p)$ has the real roots $s_j = \pm k_{Rj} = \pm p/c_{Rj}$, $j = 1,2,3$. The c_{Rj} are the three Rayleigh surface wave speeds. Speed c_{R1} is that of the well-known Rayleigh wave that propagates along a traction-free elastic surface. The *other two roots* have also been studied at length, but in boundary-initial value problems for the homogeneous elastic solid they usually do not arise and indeed are *inadmissible* in the solution. They are commonly referred to as nonphysical since they can correspond to events propagating with speeds greater than the dilatational wave speed.

It is these nonphysical roots that lead to residues that must be ruled out of the quarter-plane solutions. These residues generate *four integral equations for the time-transformed edge unknowns.* A special contour integration over Riemann surfaces, comprised of the s-plane and adjacent sheets, is needed to generate these inadmissible Rayleigh roots.

It follows that the quasi-formal solution for a problem has the form of a double inversion integral over the Bromwich paths Br_p and Br_s in the complex p and s-planes, respectively (quasi, since it contains the time-transformed edge unknowns). However, only inversion in the s-plane is needed, since, it turns out that in these problems inversion can be carried out with the Cagniard-deHoop method.

For the dilatational wavefront and Rayleigh surface wave events the forms of the time-transformed unknowns on the loaded edge $x = 0$, are known (cf. Miklowitz [11] or Rosenfeld and Miklowitz [12]), e.g.,

$$\bar{u}_d(0,y,p) \simeq A(p)e^{-k_d y}/\sqrt{y}, \quad \bar{v}_d(0,y,p) \simeq B(p)e^{-k_d y}/\sqrt{y}$$

$$\bar{u}_R(0,y,p) \simeq C(p)e^{-k_{R1} y}, \quad \bar{v}_R(0,y,p) \simeq D(p)e^{-k_{R1} y}.$$

This leaves only the four p-dependent coefficients, e.g., $A(p)$, etc., to be determined. Substitution of these forms into the four integral equations, and performing simple integrations, reduce these integral equations to *four algebraic equations* for the *coefficients* $A(p)$, $B(p)$, $C(p)$, *and* $D(p)$. Solution of these establishes the *time-transformed knowns on the loaded edge* $x = 0$. The latter are then substituted into the *quasi-formal solution* which then becomes the *formal solution*, valid for high frequency, i.e., p large. The wavefronts are then deduced by applying the technique in [11] or [12] which makes use of the Cagniard-deHoop inversion of the Br_s integral and the p-large asymptotics of the Laplace transform.

Method of Solution

In this section the present method for solving nonseparable quarter-plane problems is developed.

General Quasi-Formal Solutions. First the governing equations for the general elastodynamic quarter-plane

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

problem for plane strain will be set down. The displacement equations of motion are

$$u_{xx}(x,y,t) + (k^2-1)k^{-2}v_{xy} + k^{-2}u_{yy} = c_d^{-2}\ddot{u},$$

$$v_{xx}(x,y,t) + (k^2-1)u_{xy} + k^2 v_{yy} = c_s^{-2}\ddot{v} \tag{1}$$

for $x > 0$, $y > 0$, and time $t > 0$ where $c_d^2 = (\lambda + 2\mu)/\rho$ and $c_s^2 = \mu/\rho$ are, respectively, the squares of the dilatational and equivoluminal body wave speeds, λ and μ are Lame's constants, ρ is the material density, and $k^2 = c_d^2/c_s^2$. Corresponding stress-strain relations are

$$\sigma_x(x,y,t)/(\lambda+2\mu) = u_x(x,y,t) + (k^2-2)k^{-2}v_y(x,y,t),$$

$$\sigma_y(x,y,t)/(\lambda+2\mu) = (k^2-2)k^{-2}u_x(x,y,t) + v_y(x,y,t),$$

$$\sigma_{xy}(x,y,t)/\mu = v_x(x,y,t) + u_y(x,y,t), \quad \text{and} \quad \sigma_z = \nu(\sigma_x + \sigma_y) \tag{2}$$

where ν is Poisson's ratio. Subscripts in (1) and (2), and generally in this work, when associated with displacement, indicate partial differentiation, but when associated with stress, identify the component. Time t partial differentiation is indicated by ($\dot{}$). Initial conditions are taken as

$$u(x,y,0) = \dot{u}(x,y,0) = v(x,y,0) = \dot{v}(x,y,0) = 0 \tag{3}$$

for $x \geq 0$, $y \geq 0$, and radiation conditions as

$$\lim_{r\to\infty} \{u, u_x, v, v_x, \text{etc.}\} = 0 \tag{4}$$

for $t \geq 0$, where $r = (x^2 + y^2)^{1/2}$. Uniform loadings along an edge, and the finite amplitude plane waves they generate will, of course, not be required to satisfy (4).

We now apply one-sided Laplace transforms on t, parameter p, and x, parameter s, to (1) with the result

$$\bar{\tilde{u}}_{yy}(s,y,p) + (k^2-1)s\bar{\tilde{v}}_y(s,y,p) + (k^2s^2 - k_s^2)\bar{\tilde{u}}(s,y,p)$$
$$= k^2[s\bar{u}(0,y,p) + \bar{u}_x(0,y,p)] + (k^2-1)\bar{v}_y(0,y,p)$$
$$= f(s,y,p),$$

$$\bar{\tilde{v}}_{yy}(s,y,p) + (k^2-1)k^{-2}s\bar{\tilde{u}}_y(s,y,p) + (s^2 - k_s^2)k^{-2}\bar{\tilde{v}}(s,y,p)$$
$$= k^{-2}[s\bar{v}(0,y,p) + \bar{v}_x(0,y,p)] + (k^2-1)k^{-2}\bar{u}_y(0,y,p)$$
$$= g(s,y,p) \tag{5}$$

where a bar and a tilda over quantities indicates Laplace transforms on t and x, respectively.

The general quasi-formal solutions, based on (1), (3), and (5) are

$$\begin{Bmatrix} u(x,y,t) \\ v(x,y,t) \end{Bmatrix} = \frac{1}{2\pi i}\int_{Br_p} e^{pt}\left[\frac{1}{2\pi i}\int_{Br_s}\begin{Bmatrix}\tilde{\tilde{u}}(s,y,p)\\ \tilde{\tilde{v}}(s,y,p)\end{Bmatrix} e^{sx}\,ds\right]dp \tag{6a}$$

where

$$\tilde{\tilde{u}}(s,y,p) = [A_\alpha(s,p) - \mu_\alpha(s,y,p)]e^{-\alpha y} + [A_{-\alpha}(s,p)$$
$$+ \mu_{-\alpha}(s,y,p)]e^{\alpha y} + [A_\beta(s,p) - \mu_\beta(s,y,p)]e^{-\beta y}$$
$$+ [A_{-\beta}(s,p) + \mu_{-\beta}(s,y,p)]e^{\beta y},$$

$$\tilde{\tilde{v}}(s,y,p) = -\frac{\alpha}{s}\{[A_\alpha(s,p) - \mu_\alpha(s,y,p)]e^{-\alpha y} - [A_{-\alpha}(s,p)$$
$$+ \mu_{-\alpha}(s,y,p)]e^{\alpha y}\} + \frac{s}{\beta}\{[A_\beta(s,p) - \mu_\beta(s,y,p)]e^{-\beta y}$$
$$- [A_{-\beta}(s,p) + \mu_{-\beta}(s,y,p)]e^{\beta y}\}, \tag{6b}$$

$$\alpha = (k_d^2 - s^2)^{1/2}, \quad \beta = (k_s^2 - s^2)^{1/2}, \quad k_d^2 = p^2/c_d^2, \quad k_s^2 = p^2/c_s^2, \tag{6c}$$

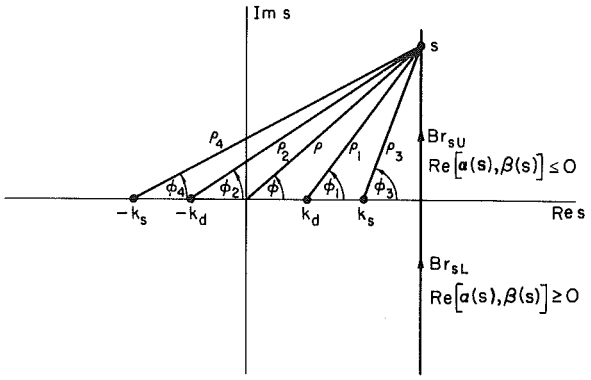


Fig. 2 Defining complex vectors for s, $s-k_d$, $s+k_d$, $s-k_s$, $s+k_s$, and $\alpha = i(s^2-k_d^2)^{1/2}$, $\beta = i(s^2-k_s^2)^{1/2}$

and where

$$\mu_{\pm\alpha} = \frac{s}{2k_s^2}\int_0^y\left[\frac{s}{\alpha}f(s,y',p) \mp k^2 g(s,y',p)\right]e^{\pm\alpha y'}\,dy',$$

$$\mu_{\pm\beta} = \frac{1}{2k_s^2}\int_0^y\left[\beta f(s,y',p) \pm k^2 s g(s,y',p)\right]e^{\pm\beta y'}\,dy'. \tag{6d}$$

The terms in (6b) associated with $A_{\pm\alpha}$ and $A_{\pm\beta}$ represent the complimentary functions of $\tilde{\tilde{u}}$ and $\tilde{\tilde{v}}$, and the $\mu_{\pm\alpha}$ and $\mu_{\pm\beta}$ terms represent the particular integrals. The solutions (6a) are called *quasi*-formal solutions since they contain time-transformed edge ($x = 0$) unknowns (displacements and displacement gradients through the f and g functions in (6d)). These unknowns do become *known* quantities for particular problems later in the method, which then makes (6a) a formal solution. Note that Br_p and Br_s are Bromwich contours in the p and s-planes, respectively.

Associated with (6a)–(6d) are the radiation conditions on y, according to (4), which can be met with suitable values involving the α and β functions in (6c). With interest in an s-plane inversion, we fix p and require that it be real and > 0. Then, as shown in Fig. 2, taking cuts to the left from branch points $s = \pm k_s, \pm k_d$, we can represent analytic branches of $\alpha(s)$ and $\beta(s)$ by

$$\alpha(s) = (k_d^2 - s^2)^{1/2} = i(s^2 - k_d^2)^{1/2} = (\rho_1\rho_2)^{1/2}e^{i(\psi+\pi/2)}, \tag{7}$$

where

$$\rho_1 = |s - k_d|, \quad \rho_2 = |s + k_d|, \quad \psi = (\varphi_1 + \varphi_2)/2$$

and

$$\beta(s) = (k_s^2 - s^2)^{1/2} = i(s^2 - k_s^2)^{1/2} = (\rho_3\rho_4)^{1/2}e^{i(\eta+\pi/2)}, \tag{8}$$

where

$$\rho_3 = |s - k_s|, \quad \rho_4 = |s + k_s|, \quad \eta = (\varphi_3 + \varphi_4)/2.$$

Hence (cf. Fig. 2), we have

$$-\frac{\pi}{2} < \begin{Bmatrix}\psi\\ \eta\end{Bmatrix} \leq 0, \ \therefore 0 < \begin{Bmatrix}\psi+\pi/2\\ \eta+\pi/2\end{Bmatrix} \leq \frac{\pi}{2} \text{ and } \therefore \text{Re}\begin{Bmatrix}\alpha(s)\\ \beta(s)\end{Bmatrix} \geq 0 \tag{9}$$

on Br_{sL}, the lower half of Br_s, and

$$0 \leq \begin{Bmatrix}\psi\\ \eta\end{Bmatrix} < \frac{\pi}{2}, \ \therefore \frac{\pi}{2} \leq \begin{Bmatrix}\psi+\pi/2\\ \eta+\pi/2\end{Bmatrix} < \pi \text{ and } \therefore \text{Re}\begin{Bmatrix}\alpha(s)\\ \beta(s)\end{Bmatrix} \leq 0 \tag{10}$$

on Br_{sU}, the upper half of Br_s.

It follows then from (9), that the multivalued functions $\alpha(s)$ and $\beta(s)$ satisfy

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\mathrm{Re}\begin{Bmatrix}\alpha(s)\\ \beta(s)\end{Bmatrix}\geq 0, \text{ for } \mathrm{Re}\{p\}>0,$$

on Br_{sL} (and Br_p), and hence, according to (6*b*), will satisfy (4) for $y\to\infty$, provided

$$A_{-\alpha(s)}=-\mu^{\infty}_{-\alpha(s)}(s,p)$$
$$=-\frac{s}{2k_s^2}\int_0^{\infty}\left[\frac{s}{\alpha(s)}f(s,y',p)+k^2g(s,y',p)\right]e^{-\alpha(s)y'}\,dy',$$

$$A_{-\beta(s)}=-\mu^{\infty}_{-\beta(s)}(s,p)$$
$$=-\frac{1}{2k_s^2}\int_0^{\infty}[\beta(s)f(s,y',p)-k^2sg(s,y',p)]e^{-\beta(s)y'}\,dy'. \tag{11}$$

Substituting (11) into (6*b*) reduces the latter to

$$\tilde{\bar{u}}(s,y,p)=[A(s,p)-\mu_{\alpha(s)}(s,y,p)]e^{-\alpha(s)y}$$
$$+[\mu_{-\alpha(s)}(s,y,p)-\mu^{\infty}_{-\alpha(s)}(s,p)]e^{\alpha(s)y}$$
$$+[B(s,p)-\mu_{\beta(s)}]e^{-\beta(s)y}+[\mu_{-\beta(s)}-\mu^{\infty}_{-\beta(s)}]e^{\beta(s)y},$$
$$\tilde{\bar{v}}(s,y,p,)=-\frac{\alpha(s)}{s}\{[A(s,p)-\mu_{\alpha(s)}]e^{-\alpha(s)y}$$
$$-[\mu_{-\alpha(s)}-\mu^{\infty}_{-\alpha(s)}]e^{\alpha(s)y}\}$$
$$+\frac{s}{\beta(s)}\{[B(s,p)-\mu_{\beta(s)}]e^{-\beta(s)y}-[\mu_{-\beta(s)}-\mu^{\infty}_{-\beta(s)}]e^{\beta(s)y}\} \tag{12}$$

where we have used $A(s,p)$ for $A_{\alpha(s)}$ and $B(s,p)$ for $A_{\beta(s)}$, and have, for simplicity, dropped the arguments for the μ's after the first line.

Similarly, from (10), the multivalued functions $\alpha(s)$ and $\beta(s)$ satisfy

$$\mathrm{Re}\begin{Bmatrix}\alpha(s)\\ \beta(s)\end{Bmatrix}\leq 0, \text{ for } \mathrm{Re}\{p\}>0,$$

on Br_{sU}(and Br_p), and hence, according to (6*b*) will satisfy (4) for $y\to\infty$, provided

$$A_{\alpha(s)}=\mu^{\infty}_{\alpha(s)}$$
$$=\frac{s}{2k_s^2}\int_0^{\infty}\left[\frac{s}{\alpha(s)}f(s,y',p)-k^2g(s,y',p)\right]e^{\alpha(s)y'}\,dy',$$

and

$$A_{\beta(s)}=\mu^{\infty}_{\beta(s)}$$
$$=\frac{1}{2k_s^2}\int_0^{\infty}[\beta(s)f(s,y',p)+k^2sg(s,y',p)]e^{\beta(s)y'}\,dy'. \tag{13}$$

Substituting (13) into (6*b*) reduces the latter to

$$\tilde{\bar{u}}(s,y,p)=[A(s,p)+\mu_{-\alpha(s)}(s,y,p)]e^{\alpha(s)y}$$
$$-[\mu_{\alpha(s)}(s,y,p)-\mu^{\infty}_{\alpha(s)}(s,p)]e^{-\alpha(s)y}$$
$$+[B(s,p)+\mu_{-\beta(s)}]e^{\beta(s)y}-[\mu_{\beta(s)}-\mu^{\infty}_{\beta(s)}]e^{-\beta(s)y},$$
$$\tilde{\bar{v}}(s,y,p)=\frac{\alpha(s)}{s}\{[A(s,p)+\mu_{-\alpha(s)}]e^{\alpha(s)y}$$
$$+[\mu_{\alpha(s)}-\mu^{\infty}_{\alpha(s)}]e^{-\alpha(s)y}\}$$
$$-\frac{s}{\beta(s)}\{[\beta(s,p)+\mu_{-\beta(s)}]e^{\beta(s)y}$$
$$+[\mu_{\beta(s)}-\mu^{\infty}_{\beta(s)}]e^{-\beta(s)y}\} \tag{14}$$

where we have used $A(s,p)$ for $A_{-\alpha(s)}$ and $B(s,p)$ for $A_{-\beta(s)}$, and again have, for simplicity, dropped the arguments for the μ's after the first line.

It is important to point out now that the branch functions $\alpha(s)$ and $\beta(s)$, defined in (7) and (8), respectively, are analytic all along Br_s, having

$$\mathrm{Re}\begin{bmatrix}\alpha(s)\\ \beta(s)\end{bmatrix}\geq 0 \text{ on } Br_{sL} \text{ from (9) and } \mathrm{Re}\begin{bmatrix}\alpha(s)\\ \beta(s)\end{bmatrix}\leq 0$$

on Br_{sU} from (10). Equations (12) for $\tilde{\bar{u}}$ and $\tilde{\bar{v}}$ are bounded as $y\to\infty$, the critical terms being those involving $\exp[\alpha(s)y]$ and $\exp[\beta(s)y]$ where

$$\mathrm{Re}\begin{bmatrix}\alpha(s)\\ \beta(s)\end{bmatrix}\geq 0$$

and the coefficients of these exponential functions vanish as $y\to\infty$. Similarly, (14) for $\tilde{\bar{u}}$ and $\tilde{\bar{v}}$ are bounded as $y\to\infty$, since the critical terms being those involving $\exp[-\alpha(s)y]$ and $\exp[-\beta(s)y]$ where

$$\mathrm{Re}\begin{bmatrix}\alpha(s)\\ \beta(s)\end{bmatrix}\leq 0$$

and the coefficients of these exponential functions vanish as $y\to\infty$. It is of further interest to note from the results of (9) and (10) that according to the principle of reflection $\alpha(s)=-\overline{\alpha(s)}$ and $\beta(s)=-\overline{\beta(s)}$ on Br_{sU}, so that (14) forms a natural conjugate of (12). This leads to a real formal solution for the displacements over Br_{sL}, as well as other conjugations in the s-plane.

The Pressure Shock Problem. The pressure shock problem, shown in Fig. 1, is now introduced. The double transformed boundary conditions for this problem, obtained from (2), are for $y=0$,

$$\tilde{\bar{\sigma}}_y(s,0,p)/(\lambda+2\mu)=(k^2-2)k^{-2}[s\tilde{\bar{u}}(s,0,p)$$
$$-\bar{u}(0,0,p)]+\tilde{\bar{v}}_y(s,0,p)=0,$$
$$\tilde{\bar{\sigma}}_{yx}(s,0,p)/\mu=\tilde{\bar{u}}_y(s,0,p)+s\tilde{\bar{v}}(s,0,p)-\bar{v}(0,0,p)=0, \tag{15}$$

and for $x=0$,

$$\bar{\sigma}_x(0,y,p)/(\lambda+2\mu)=\bar{u}_x(0,y,p)$$
$$+(k^2-2)k^{-2}\bar{v}_y(0,y,p)=\sigma_0/(\lambda+2\mu)p,$$
$$\bar{\sigma}_{xy}(0,y,p)/\mu=\bar{v}_x(0,y,p)+\bar{u}_y(0,y,p)=0, \tag{16}$$

where we have taken $f(t)=H(t)$ the Heaviside step. Now from (12) we derive $\tilde{\bar{u}}_y(s,y,p)$ and $\tilde{\bar{v}}_y(s,y,p)$, and substituting these and (12) into (15), we find $A(s,p)$ and $B(s,p)$ to be

$$A(s,p)=A_N(s,p)/R(s,p)=\begin{vmatrix}q(s,p) & -2s^2\\ r(s,p) & k_s^2-2s^2\end{vmatrix}/R(s,p),$$
$$B(s,p)=B_N(s,p)/R(s,p)=\begin{vmatrix}k_s^2-2s^2 & q(s,p)\\ 2\alpha\beta & r(s,p)\end{vmatrix}/R(s,p), \tag{17}$$

where

$$q(s,p)=(k_s^2-2s^2)\mu^{\infty}_{-\alpha}-2s^2\mu^{\infty}_{-\beta}+(k^2-2)s\bar{u}(0,0,p),$$
$$r(s,p)=-[2\alpha\beta\mu^{\infty}_{-\alpha}+(k_s^2-2s^2)\mu^{\infty}_{-\beta}+\beta\bar{v}(0,0,p)],$$

and

$$R(s,p)=(k_s^2-2s^2)^2+4s^2\alpha\beta$$

is the generalized Rayleigh function. Further, using (16), in $f(s,y,p)$ and $g(s,y,p)$ the latter reduce to

$$f(s,y,p)=k^2s\bar{u}(0,y,p)+\bar{v}_y(0,y,p)+\sigma_0/\mu p,$$
$$g(s,y,p)=k^{-2}[s\bar{v}(0,y,p)+(k^2-2)\bar{u}_y(0,y,p)]. \tag{18}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

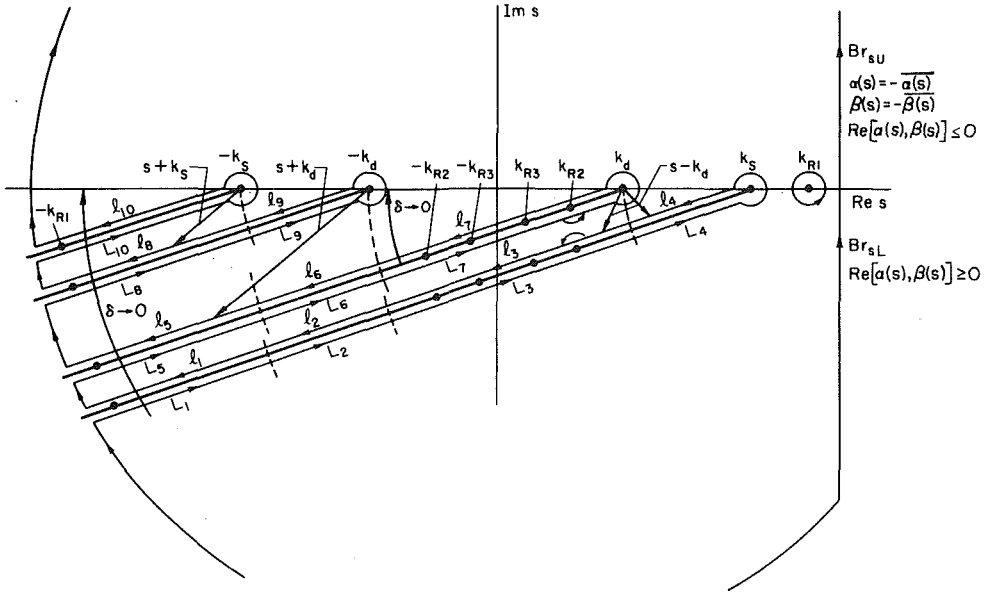


Fig. 3 Limiting nest of contours as $\delta \to 0$

Substituting (17) and (18) into (12) defines the doubly transformed displacement solutions $\tilde{\tilde{u}}(s,y,p)$ and $\tilde{\tilde{v}}(s,y,p)$, from Br_{sL}. Similarly, from (14) we have

$$\tilde{\tilde{u}}(\bar{s},y,p) = [A(\bar{s},p) + \mu_{\overline{\alpha(s)}}]e^{-\overline{\alpha(s)}y} - [\mu_{-\overline{\alpha(s)}} - \mu^{\infty}_{-\overline{\alpha(s)}}]e^{\overline{\alpha(s)}y} + [B(\bar{s},p) + \mu_{\overline{\beta(s)}}]e^{-\overline{\beta(s)}y} - [\mu_{-\overline{\beta(s)}} - \mu^{\infty}_{-\overline{\beta(s)}}]e^{\overline{\beta(s)}y},$$

$$\tilde{\tilde{v}}(\bar{s},y,p) = -\frac{\overline{\alpha(s)}}{\bar{s}}\{[A(\bar{s},p) + \mu_{\overline{\alpha(s)}}]e^{-\overline{\alpha(s)}y} + [\mu_{-\overline{\alpha(s)}} - \mu^{\infty}_{-\overline{\alpha(s)}}]e^{\overline{\alpha(s)}y}\} + \frac{\bar{s}}{\overline{\beta(s)}}\{[B(\bar{s},p) + \mu_{\overline{\beta(s)}}]e^{-\overline{\beta(s)}y} + [\mu_{-\overline{\beta(s)}} - \mu^{\infty}_{-\overline{\beta(s)}}]e^{\overline{\beta(s)}y}\} \tag{19}$$

where

$$A(\bar{s},p) = A_N(\bar{s},p)/R(\bar{s},p) = \begin{vmatrix} q(\bar{s},p) & -2\bar{s}^2 \\ r(\bar{s},p) & k^2 - 2\bar{s}^2 \end{vmatrix} / R(\bar{s},p),$$

$$B(\bar{s},p) = B_N(\bar{s},p)/R(\bar{s},p) = \begin{vmatrix} k_s^2 - 2\bar{s}^2 & q(\bar{s},p) \\ 2\alpha(s)\ \beta(s) & r(\bar{s},p) \end{vmatrix} / R(\bar{s},p), \tag{20}$$

where

$$q(\bar{s},p) = -(k_s^2 - 2\bar{s}^2)\mu^{\infty}_{-\overline{\alpha(s)}} + 2\bar{s}^2\mu^{\infty}_{-\overline{\beta(s)}} + (k^2 - 2)\bar{s}\bar{u}(0,0,p),$$

$$r(\bar{s},p) = 2\overline{\alpha(s)}\ \overline{\beta(s)}\,\mu^{\infty}_{-\overline{\alpha(s)}} + (k_s^2 - 2\bar{s}^2)\mu^{\infty}_{-\overline{\beta(s)}} - \overline{\beta(s)}\,\bar{v}(0,0,p),$$

$$R(\bar{s},p) = (k_s^2 - 2\bar{s}^2)^2 + 4\bar{s}^2\overline{\alpha(s)}\ \overline{\beta(s)},$$

where it should be noted, in comparing (19) and (20) with (12) and (17), that important changes from $\alpha(s)$ and $\beta(s)$ to $\overline{\alpha(s)}$ and $\overline{\beta(s)}$ has resulted in the exponential functions and the subscripts of all the μ-functions, the latter functions having a sign change as well. Equations (19) and (20) together with (12), (17), and (18) lead to the formal inversion through (6*a*), as discussed earlier.

Indicated Contour Integration for Solution. As pointed out in the Introduction one needs a somewhat special contour integration in the *s*-plane to solve the present class of quarter-plane problems, one in which the so-called nonphysical roots and associated poles and residues of the generalized Rayleigh function are generated and, hence, can be exploited in the form of conditions on the solution ruling out unbounded and inadmissible events. A general contour that can do this is depicted in Fig. 3. It is composed of a nest of component contours, based on four branch cuts taken to the left of the branch points, $s = \pm k_d$ and $\pm k_s$. It may be seen that the cuts, and contours along them, are in a rotated (by a small counterclockwise angle δ) position below the real axis. Indeed the separation of the cuts and related contours in this nest were convenient in keeping track of the vectors $s \mp k_d$, $s \mp k_s$ as the contours were negotiated. More important, however, was the fact that this nest of contours did its job of generating the nonphysical zeros (and corresponding poles) of the generalized Rayleigh function. As the large arrow in the clockwise direction across the cuts and contours in Fig. 3 indicates, these cuts and contours are limiting ones that, as $\delta \to 0$, move to superposed positions along the real axis.

Certain other features of the cuts and contours in Fig. 3 should be pointed out. It turns out that oppositely directed lineal path contributions along *inside* contours cancel, e.g., the whole or parts of the paths L_5, L_6, and L_7 and their mates l_1, l_2, and l_3, respectively. To show that this is the case we proceed as follows. Consider first, integration along the first contour (in Fig. 3) L_1, L_2, L_3, L_4, and l_4. It is straightforward with the four vectors $s \pm k_s$, $s \pm k_d$ in expected positions. Now, however, $s - k_d$ leaves path l_4 to begin its excursion along paths l_3, l_2, and l_1. Note though, in the final position of these contours (i.e., when $\delta \to 0$), they will be just an ϵ above the real axis. This position for l_3, l_2, and l_1 has taken $s - k_d$ across its cut to the beginning of the next sheet[1], $\arg(s - k_d) = -(\pi + \epsilon)$, of the Riemann surface for the function $(s - k_d)^{1/2}$. We see then that from the end of path l_1, $s - k_d$ can rotate counterclockwise a distance 2ϵ back to the first sheet and begin its negotiation of paths L_5, L_6, L_7, l_7, l_6, and l_5. These paths all lie in the first sheet. Such a process gives

[1] The *s*-plane in Fig. 3 can be taken as the first sheet $-\pi < \arg(s - k_d) < \pi$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

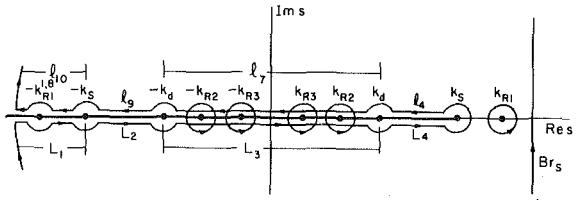


Fig. 4 Remaining contours after $\delta \to 0$

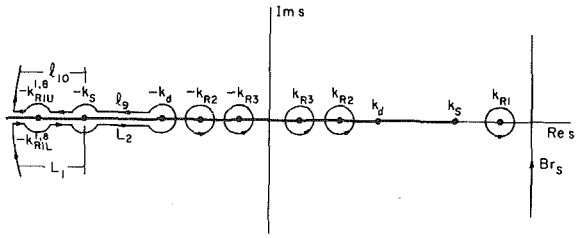


Fig. 5 Remaining contours after imposing solution boundedness and wave speed restriction ($c \leq c_d$) to path integrals L_3, L_4, l_4, and l_7

cancellation of the oppositely directed lineal path contributions from the respective pairs of the whole or parts of the paths L_5, L_6, L_7, and l_1, l_2, l_3.

It should be noted from Fig. 3 that a similar process involves the complex vectors $s + k_d$, $s + k_s$, their cuts and related contours. The complex vector $s - k_s$ does not leave the first sheet of $(s - k_s)^{1/2}$.

As noted earlier this integration generates the nonphysical, as well as the physical, zeros of the generalized Rayleigh function. These zeros are $s = s_{\pm j}(p) = \pm k_{Rj} = \pm p/\gamma_j c_s$, γ_j being positive constants[2] with $j = 1, 2, 3$. The nonphysical roots are $s = s_{\pm j}(p) = \pm k_{Rj}$, $j = 2, 3$, and as Fig. 3 shows (assuming $\delta \to 0$), lie along the real axis between $-k_d$ and k_d. As the figure indicates these zeros lie on the line $\arg(s - k_d) = -\pi$, common to both the first and second sheets of the Riemann surface. It follows that the poles associated with these zeros generate the half-circular paths, indicated in Fig. 3 for the $s_2(p) = k_{R2}$ case, along l_3 and L_7. We see the half circle along L_7 lies on the ending of the first sheet, and the half circle along l_3 on the beginning of the second sheet. Hence, their sum is a complete continuous counterclockwise circular path [about zero $s_2(p) = k_{R2}$] lying on the Riemann surface. The other nonphysical Rayleigh function zeros generate similar circular paths in the contour integration.

On the basis of the foregoing considerations the contour integration reduces to that shown in Fig. 4. The contributing contours are just L_1, L_2, L_3, L_4, and l_4, l_7, l_9, l_{10} plus six circular paths for the poles associated with the Rayleigh function roots $s = s_{\pm j}(p) = \pm k_{Rj}$, $j = 1, 2, 3$. The paths around $s_{\pm j}(p) = \pm k_{Rj}$, $j = 2, 3$ are all free of the cuts, and lineal paths intersecting them, which is indicated in Fig. 4 by complete circles. Circular paths are also shown for $s_{\pm 1}(p) = \pm k_{R1}$. As Fig. 3 shows, there must be two such paths for $-k_{R1}$, associated with the pairs of paths L_1 and l_{10} and L_8 and l_5.

The radiation condition (4) reduces the number of contours in Fig. 4 still further. Figure 4 (and the analysis leading to it) show that the right half s-plane linear path integrals, parts of L_3 and l_7 for $0 < s < k_d$, and the whole of L_4 and l_4 for $k_d < s < k_s$, all represent exponentially unbounded contributions through the kernel of the Br_s integral e^{sx}, for $x > 0$ and Re $p > 0$. The radiation condition (4) therefore rules out these contributions. They can also be identified with waves traveling in the negative x-direction by noting the double inversion integral kernel $\exp\{p[t + (s/p)x]\}$ would then be $\exp\{p[t - (x/-c)]\}$, since s has the dimensions of p/c where c is the wave speed. Equation (4) would also rule out these waves. Further, consider the parts of L_3 and l_7 integrals in the left-half s-plane, where $-k_d < s < 0$, also having the double inversion integral kernel $\exp\{p[t + (s/p)x]\}$. The s interval here may be written as $-k_d < s = p/c < 0$ which corresponds to $c_d < c < \infty$, hence $\exp\{p[t + (s/p)x]\} = \exp\{p[t - (x/c)]\}$ for these integrals. These would be waves propagating in the positive x-direction, but they have speed $c > c_d$ and hence are inadmissible in a linear elastic solid. We are thus left with the contours depicted in Fig. 5. It is important to point out that this analysis, ruling out unbounded contributions and inadmissible events, is related to the procedure usually used in separable elastodynamic wave problems. It is, of course, much simpler in the latter, usually done early in the analysis by setting the coefficients of the two unbounded kernels, like $e^{\alpha y}$ and $e^{\beta y}$ here, equal to zero. The related procedure here is more complicated and has come along later in analysis.

Consistently, we now require that the contributions (residues) from the nonphysical poles, generated by the zeros of the generalized Rayleigh function, $s_{\pm j}(p) = \pm k_{Rj}$, $j = 2, 3$, and $s_1(p) = k_{R1}$, must be ruled out of the solution ($k_{Rj} = p/c_{Rj}$, where c_{Rj} is the wave speed). Of these, the residues at the poles $s_j(p) = k_{Rj}$, $j = 1, 2, 3$ are ruled out by the boundedness condition on the solution. The residues at $s_{-j}(p) = -k_{Rj}$, $j = 2, 3$, are ruled out because the corresponding speed of these Rayleigh wave events are greater than c_d. Hence, as the next section points out the residues for these five poles are set equal to zero. This eliminates the circular paths in Fig. 5 for these contributions, generating four coupled integral equations that guarantee boundedness for the solution and the elimination of inadmissible events. Solution of these integral equations determines the edge unknowns in our problem, at least for wavefront (high frequency) events.

Conditions for Solution Boundedness, Inadmissible Events. Derivation of the four integral equations for the edge ($x = 0$) unknowns begins with first establishing that $s_{\pm j}(p) = \pm k_{Rj}$, $j = 2, 3$, and $s_1(p) = k_{R1}$ are zeros of the generalized Rayleigh function. Consider for example, $s_2(p) = k_{R2} = p/\gamma_2 c_s$ which lies along l_3, L_7 in Fig. 3. Using the limiting values of s, $s \pm k_d$, $s \pm k_s$, α and β on these paths, with Poisson's ratio 1/4, hence $\gamma_2 \simeq 1.78$, one finds by substituting these values in $c_s^4 R(s,p)/p^4$ that the latter will vanish since

$$
\begin{aligned}
R(s,p)\,|_{s=k_{R2}} &= [(k_s^2 - 2s^2)^2 + 4s^2\alpha(s)\beta(s)]_{s=k_{R2}} \\
&= [k_s^2 - 2s^2)^2 - 4s^2\alpha(s)\overline{\beta(s)}]_{s=k_{R2}} \\
&= [(k_s^2 - 2s^2)^2 - 4s^2(\rho_1\rho_2)^{1/2}(\rho_3\rho_4)^{1/2}]_{s=k_{R2}} \\
&= (k_s^2 - 2k_{R2}^2)^2 - 4k_{R2}^2(k_d^2 - k_{R2}^2)^{1/2}(k_s^2 - k_{R2}^2)^{1/2} \qquad (21)
\end{aligned}
$$

vanishes. Note that ρ_1, ρ_2, ρ_3 and ρ_4, defined generally in (7) and (8), are at $s = k_{R2}$, $k_d - k_{R2}$, $k_d + k_{R2}$, $k_s - k_{R2}$, and $k_s + k_{R2}$, respectively. Note also $\alpha(s) = (\rho_1\rho_2)^{1/2}$ and $\beta(s) = (\rho_3\rho_4)^{1/2} = -\overline{\beta(s)}$ in (21) are continuations from Br_{sL} and Br_{sU}, respectively, and that these forms hold for $s = -k_{R2}$, $\pm k_{R3}$ as well. Similarly one can establish that $s = k_{R1}$, and the admissible $s = -k_{R1}$ are zeros.

The residues at $s = k_{R1}$, $\pm k_{Rj}$, $j = 2, 3$ can be shown to be nonvanishing through the usual calculations.[3] These residues can then be set equal to zero by requiring that

[2] The γ_j's depend on Poisson's ratio. For Poisson's ratio = 1/4, $\gamma_1 \simeq 0.92$, $\gamma_2 \simeq 1.78$, and $\gamma_3 = 2$. These are the γ_j's we have chosen for our problem.

[3] These calculations are like those in the later section dealing with the physical Rayleigh waves on the free edge $y = 0$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$A_N(s,p)\,|_{s=k_{R1},\pm k_{Rj}} = B_N(s,p)\,|_{s=k_{R1},\pm k_{Rj}} = 0,\ j=2,3, \qquad (22)$$

where $A_n(s,p)$, $B_N(s,p)$ are given in (17). Now with $s_{\pm j}(p) = \pm k_{Rj}$, $j = 2, 3$ and $s_1(p) = k_{R1}$ representing the five Rayleigh poles here, the two equations (22) are expanded using (17) and (6*d*). After some algebra, and a simple integration on y for the coefficient of the input term $\sigma_0/\mu p$, the general integral equation is found to be

$$\int_0^\infty \left\{ \frac{s_j}{\alpha_j}[(k_s^2-2s_j^2)e^{-\alpha_j y} - 2\alpha_j\beta_j e^{-\beta_j y}] \right.$$
$$[k^2 s_j \bar{u}(0,y,p) + \bar{v}_y(0,y,p)]$$
$$+ [(k_s^2-2s_j^2)e^{-\alpha_j y} + 2s_j^2 e^{-\beta_j y}]$$
$$\left.[s_j\bar{v}(0,y,p) + (k^2-2)\bar{u}_y(0,y,p)]\right\} dy$$
$$+(k^2-2)k_s^2\bar{u}(0,0,p) + \frac{k_s^2(k_s^2-2s_j^2)}{2s_j\alpha_j}\bar{v}(0,0,p)$$
$$= \left(2 - \frac{k_s^2-2s_j^2}{\alpha_j^2}\right)\frac{s_j\sigma_0}{\mu p}, \qquad (23)$$

where $s_j = s_j(p)$, $j = 1, \pm 2, \pm 3$. The s_j are basic to the definitions of the α_j and β_j.

Consider α_1 and β_1 first. From (7) and (8) we have, when $s_1 = k_{R1}$ (cf. Fig. 3), $\alpha_1 = i(\rho_1\rho_2)^{1/2} = i(k_{R1}^2 - k_d^2)^{1/2}$ and $\beta_1 = i(\rho_3\rho_4)^{1/2} = i(k_{R1}^2 - k_s^2)^{1/2}$. Likewise, when $s_{\pm 2} = \pm k_{R2}$, $s_{\pm 3} = \pm k_{R3}$, along paths l_3 and L_7, (7) and (8) show that $\alpha_{\pm 2} = (\rho_1\rho_2)^{1/2} = (k_d^2 - k_{R2}^2)^{1/2}$ and $\beta_{\pm 2} = -(\rho_3\rho_4)^{1/2} = -\bar{\beta}_{\pm 2} = -(k_s^2 - k_{R2}^2)^{1/2}$ with $\alpha_{\pm 3}$ and $\beta_{\pm 3}$ having the same forms as $\alpha_{\pm 2}$ and $\beta_{\pm 2}$ except that k_{R3} replaces k_{R2}. Equation (23) represent five coupled integral equations for the two time-transformed edge displacement unknowns $\bar{u}(0,y,p)$, $\bar{v}(0,y,p)$ and their corner values $\bar{u}(0,0,p)$, $\bar{v}(0,0,p)$.

Now substituting the $j = 1$ terms, s_1, α_1, and β_1 in the foregoing, into (23), one finds it splits into two equations, the real and imaginary parts created by the imaginary nature of α_1 and β_1. These two real integral equations provide the definition of the time-transformed corner displacements $\bar{u}(0,0,p)$ and $\bar{v}(0,0,p)$ given by

$$\begin{bmatrix} \bar{u}(0,0,p) - P(p)K(p) \\ \bar{v}(0,0,p) \end{bmatrix} = \begin{pmatrix} -P(p) \\ Q(p) \end{pmatrix} \int_0^\infty$$
$$\left\{ \mp k_{R1}\left\{ [(k_s^2-2k_{R1}^2)/|\alpha_1|] \begin{bmatrix} \sin|\alpha_1|y \\ \cos|\alpha_1|y \end{bmatrix} \right.\right.$$
$$\left.\left. + 2|\beta_1| \begin{bmatrix} \sin|\beta_1|y \\ \cos|\beta_1|y \end{bmatrix} \right\} \{k^2 k_{R1}\bar{u}(0,y,p) + \bar{v}_y(0,y,p)\}$$
$$+ \left\{ (k_s^2-2k_{R1}^2) \begin{bmatrix} \cos|\alpha_1|y \\ \sin|\alpha_1|y \end{bmatrix} + 2k_{R1}^2 \begin{bmatrix} \cos|\beta_1|y \\ \sin|\beta_1|y \end{bmatrix} \right\}$$
$$\left.\{k_{R1}\bar{v}(0,y,p) + (k^2-2)\bar{u}_y(0,y,p)\}\right\} dy \qquad (24)$$

where

$$K(p) = \{2 - [(k_s^2-2k_{R1}^2)/|\alpha_1|^2]\}\sigma_0 k_{R1}/\mu p,$$
$$P(p) = (k^2-2)^{-1}k_s^{-2}$$

and

$$Q(p) = -2k_{R1}|\alpha_1|/k_s^2(k_s^2-2k_{R1}^2).$$

Now, if these integral expressions for $\bar{u}(0,0,p)$ and $\bar{v}(0,0,p)$ are substituted into (23), then there are just four integral equations involving only the time-transformed edge unknowns $\bar{u}(0,y,p)$, $\bar{v}(0,y,p)$, and related derivatives $\bar{u}_y(0,y,p)$ and $\bar{v}_y(0,y,p)$. These four remaining equations are parametric in $s_{\pm 2}$ and $s_{\pm 3}$, respectively, with related $\alpha_{\pm 2}$, $\beta_{\pm 2}$, $\alpha_{\pm 3}$ and $\beta_{\pm 3}$. This solution provides the information required to reduce the quasi-formal solution to the formal solution which can then be inverted by known exact and approximate solution techniques. The next section shows how the four integral equations can be used to find these time-transformed edge unknowns for high-frequency events (dilatational wavefronts and Rayleigh surface waves) on the loaded edge $x = 0$ of this problem.

Determination of Time-Transformed Edge ($x = 0$) and Corner ($x = y = 0$) Unknowns for Wavefront (High-Frequency) Events and Their Inverses. To derive the time-transformed edge ($x = 0$) unknowns representing wavefront (high-frequency events) we will use as a guide similar general features of related events from the elastic half plane and waveguide problems. Indeed, we can construct forms for the time-transformed edge unknowns that will reduce the four integral equations to algebraic equations for the time-transformed p-dependent coefficients of these unknowns.

From the wavefront analysis work [7] on the plane strain semi-infinite plate, involving mixed-pressure edge conditions, we assume at the dilatational wavefront, the time-transformed displacements $\bar{u}(0,y,p)$, $\bar{v}(0,y,p)$ are, respectively,

$$\bar{u}_d(0,y,p) \simeq A(p)y^{-1/2}e^{-k_d y},$$
$$\bar{v}_d(0,y,p) \simeq B(p)y^{-1/2}e^{-k_d y}, \qquad (25)$$

from which we also have

$$\bar{u}_{yd}(0,y,p) \simeq -k_d A(p)y^{-1/2}e^{-k_d y},$$
$$\bar{v}_y(0,y,p) \simeq -k_d B(p)y^{-1/2}e^{-k_d y}. \qquad (26)$$

It may be noted the $y^{-1/2}$ in these expressions is in agreement with the classical form for the propagation of two-dimensional surfaces of discontinuity in elastodynamics. The $\exp(-k_d y)$ through its shift operator nature guarantees the expected one-dimensional wave nature of these quantities, e.g., $u_d(0,y,p) \Rightarrow u_d(0,y,t-y/c_d)$.

Consider next the Rayleigh wave disturbance on the edge $x = 0$. Here as in the half-plane problem one would expect a nondecaying in space (y) one-dimensional disturbance. Hence, it is assumed that

$$\bar{u}_R(0,y,p) = C(p)\exp(-k_{R1}y),$$
$$\bar{v}_R(0,y,p) = D(p)\exp(-k_{R1}y), \qquad (27)$$

from which one also has

$$\bar{u}_{yR}(0,y,p) = -k_{R1}C(p)\exp(-k_{R1}y),$$
$$\bar{v}_{yR}(0,y,p) = -k_{R1}D(p)\exp(-k_{R1}y). \qquad (28)$$

It follows then, neglecting the equivoluminal disturbance on the edge ($x=0$) for the moment, that the right-hand sides of the expressions (25)–(28) may be substituted into the four integral equations obtained from (23) and (24), as discussed after (24). The simple integrations these terms provide reduce the four integral equations to four algebraic ones that determine the time-transformed coefficients $A(p)$, $B(p)$, $C(p)$, and $D(p)$ through Cramer's rule. They are given by

$$A(p) = Ap^{-5/2}, \quad B(p) = Bp^{-5/2},$$
$$C(p) = Cp^{-2}, \quad D(p) = Dp^{-2} \qquad (29)$$

where A, B, C, and D are constants given in (*A*.1) of the Appendix. Substituting the right-hand sides of (29) into (25) and (27) and then inverting the latter, one finds the dilatational displacement wavefronts on the loaded edge ($x = 0$) to be

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\begin{bmatrix} u_d(0,y,t) \\ v_d(0,y,t) \end{bmatrix} = \frac{4}{3\sqrt{\pi}} \begin{bmatrix} A \\ B \end{bmatrix} y^{-1/2} (t-y/c_d)^{3/2} H(t-y/c_d), \quad (30)$$

and the Rayleigh displacement wavefronts there to be

$$\begin{bmatrix} u_R(0,y,t) \\ v_R(0,y,t) \end{bmatrix} = \begin{bmatrix} C \\ D \end{bmatrix} (t-y/c_{R1}) H(t-y/c_{R1}). \quad (31)$$

Note that the corresponding wavefronts for the velocities $\dot{u}_d$, $\dot{v}_d$, and $\dot{u}_R$, $\dot{v}_R$ behave as $(t-y/c_d)^{1/2}$ and the step $H(t-y/c_{R1})$, respectively. Accordingly, the accelerations behave as $(t-y/c_d)^{-1/2}$ and $\delta(t-y/c_{R1})$, both singular.

Concerning the two-sided equivoluminal wavefronts on the edge, $x = 0$, the present method yields them in terms of the coefficients A, B, C, and D as we shall see in Part 2. In effect this says the dilatational and Rayleigh wavefronts (on $x = 0$) generate these equivoluminal wavefronts on the edge $x = 0$, as well as the other wavefronts in the problem. Interesting is the fact that the two-sided equivoluminal wavefronts on edge $x = 0$ have the same time behavior as the dilatational wavefronts there, which is also true in Lamb's plane strain problem, cf. [11], equations (6.52), (6.53)].

To determine the corner displacements $u(0,0,t)$ and $v(0,0,t)$ the right-hand sides of (25)–(28) are substituted into (24), resulting in real integrals of the types

$$\int_0^\infty \begin{Bmatrix} \sin m\,y \\ \cos m\,y \end{Bmatrix} \frac{e^{-ay}}{\sqrt{y}}\,dy = \left[\frac{\pi}{2(a^2+m^2)}\right]^{1/2} \begin{bmatrix} [(a^2+m^2)^{1/2}-a]^{1/2} \\ [(a^2+m^2)^{1/2}+a]^{1/2} \end{bmatrix},$$

$$\int_0^\infty \begin{Bmatrix} \sin m\,y \\ \cos m\,y \end{Bmatrix} e^{-ay}\,dy = \begin{Bmatrix} m \\ a \end{Bmatrix} \Big/ (a^2+m^2), \quad (32)$$

where a and m are > 0. The right-hand sides of (32) therefore reduce (24) to algebraic equations for $\bar{u}(0,0,p)$ and $\bar{v}(0,0,p)$ involving $A(p)$, $B(p)$, $C(p)$, and $D(p)$ which, using the right-hand sides of (27) yield

$$\bar{u}(0,0,p) = \hat{u}p^{-2}, \quad \bar{v}(0,0,p) = \hat{v}p^{-2}, \quad (33)$$

where $\hat{u}$ and $\hat{v}$ are given in (A.2) of the Appendix. It follows that the corner displacements are

$$u(0,0,t) = \hat{u}t, \quad v(0,0,t) = \hat{v}t \quad (34)$$

with corresponding velocities and accelerations being time steps and delta function singularities, respectively.

The Formal Solution for the Time-Transformed Displacements $\bar{u}(x,y,p)$, $\bar{v}(x,y,p)$. With the time-transformed edge unknowns in (25)–(28) now determined by (29) and (A.1), and similarly the time-transformed corner displacements by (33) and (A.2), the quasi-formal solutions for the displacements $u(x,y,t)$, $v(x,y,t)$ [(6a–d)] become the formal solutions for high-frequency (wavefront) events for these displacements. It follows, using (7)–(10), (12), (17), (19)–(20), and (33), that these formal solutions can be written as

$$\begin{Bmatrix} \bar{u}(x,y,p) \\ \bar{v}(x,y,p) \end{Bmatrix} = \frac{1}{2\pi i}\left[\int_{Br_{sL}} e^{sx}\left\{\begin{Bmatrix} 1 \\ -\alpha/s \end{Bmatrix}\left\{[U_{dR}(s,p)-\mu_\alpha(s,y,p)]e^{-\alpha y} + [\pm]\left[\mu_{-\alpha}-\mu^\infty_{-\alpha}\right]e^{\alpha y}\right\} + \begin{Bmatrix} 1 \\ s/\beta \end{Bmatrix}\left\{[U_{sR}(s,p)-\mu_\beta(s,y,p)]e^{-\beta y} + [\pm]\left[\mu_{-\beta}-\mu^\infty_{-\beta}\right]e^{\beta y}\right\}\right\} ds + \int_{Br_{sU}} e^{\bar{s}x}\left\{\begin{Bmatrix} 1 \\ -\bar{\alpha}/\bar{s} \end{Bmatrix}\left\{[U_{dR}(\bar{s},p)+\mu_{\bar{\alpha}}(\bar{s},y,p)]e^{-\bar{\alpha}y} + [\pm]\left[\mu_{-\bar{\alpha}}-\mu^\infty_{-\bar{\alpha}}\right]e^{\bar{\alpha}y}\right\} + \begin{Bmatrix} 1 \\ \bar{s}/\bar{\beta} \end{Bmatrix}\left\{[U_{sR}(\bar{s},p)+\mu_{\bar{\beta}}(\bar{s},y,p)]e^{-\bar{\beta}y} + [\mp]\left[\mu_{-\bar{\beta}}-\mu^\infty_{-\bar{\beta}}\right]e^{\bar{\beta}y}\right\}\right\} d\bar{s}\right], \quad (35a)$$

where

$$R(s,p)\,U_{dR}(s,p) = [(k_s^2-2s^2)^2 - 4s^2\alpha\beta]\mu^\infty_{-\alpha} - 4s^2(k_s^2-2s^2)\mu^\infty_{-\beta} + s[(k^2-2)(k_s^2-2s^2)\hat{u} - 2s\beta\hat{v}]p^{-2},$$

$$R(s,p)\,U_{sR}(s,p) = -4\alpha\beta(k_s^2-2s^2)\mu^\infty_{-\alpha} - [(k_s^2-2s^2)^2 - 4s^2\alpha\beta]\mu^\infty_{-\beta} - \beta[2(k^2-2)s\alpha\bar{u} + (k_s^2-2s^2)\hat{v}]p^{-2},$$

$$R(\bar{s},p)\,U_{dR}(\bar{s},p) = -[(k_s^2-2\bar{s}^2)^2 - 4\bar{s}^2\bar{\alpha}\bar{\beta}]\mu^\infty_{-\bar{\alpha}} + 4\bar{s}^2(k_s^2-2\bar{s}^2)\mu^\infty_{-\bar{\beta}} + \bar{s}\,[(k^2-2)(k_s^2-2\bar{s}^2)\hat{u} - 2\bar{s}\bar{\beta}\hat{v}]p^{-2},$$

$$R(\bar{s},p)\,U_{sR}(\bar{s},p) = 4\bar{\alpha}\bar{\beta}(k_s-2\bar{s}^2)\mu^\infty_{-\bar{\alpha}} + [(k_s^2-2\bar{s}^2)^2 - 4\bar{s}^2\bar{\alpha}\bar{\beta}]\mu^\infty_{-\bar{\beta}} - \bar{\beta}[2(k^2-2)\bar{s}\bar{\alpha}\hat{u} + (k_s^2-2\bar{s}^2)\hat{v}]p^{-2}, \quad (35b)$$

and where the arguments (s) of α, $\bar{\alpha}$, β, and $\bar{\beta}$ have been deleted for simplicity, and $R(s,p)$, $R(\bar{s},p)$ and the μ-functions are given in (17), (12), and (6d), respectively. The subscripts d and s indicate essentially terms associated with dilatational and equivoluminal motions, respectively, and subscript R indicates association with the Rayleigh functions $R(s,p)$ and $R(\bar{s},p)$.

With the aid of the principle of reflection in complex variables, (35) can be reduced to real equations for $\bar{u}$ and $\bar{v}$ through complex conjugation. Consider, for example, the conjugation of the first terms on the right-hand sides of $U_{dR}(s,p)$ and $U_{dR}(\bar{s},p)$ in (35b), respectively, on Br_{sL} and Br_{sU} in (35a). Analysis of these terms at the intersection point of Br_{sU} and the real s-axis, $s = \gamma$, shows that

$$\begin{bmatrix} U_{dR}(s,p) \\ U_{dR}(\bar{s},p) \end{bmatrix}_{s=\gamma} = \pm A_1(\gamma,p) + iA_2(\gamma,p) = \begin{bmatrix} A(s,p) \\ -A(\bar{s},p) \end{bmatrix}_{s=\gamma} \quad (36)$$

where $\pm A(\gamma,p)$ are complex numbers with A_1 and A_2 being real. It follows from (36) by analytical continuation away from $s = \gamma$ that

$$\begin{bmatrix} U_{dR}(s,p) \\ U_{dR}(\bar{s},p) \end{bmatrix} = \begin{bmatrix} A(s,p) \\ -A(\bar{s},p) \end{bmatrix} = \begin{bmatrix} A_1(s,p)+iA_2(s,p) \\ -A_1(\bar{s},p)+iA_2(\bar{s},p) \end{bmatrix} \quad (37)$$

along Br_{sL} and Br_{sU}, respectively.

Equations (36) show that both cases of the principle of reflection are involved since both a real and an imaginary number make up $\pm A(\gamma,p)$. According to (37) the first case has the form

$$\frac{1}{2\pi i}\left[\int_{Br_{sL}} A_1(s,p)\,ds - \int_{Br_{sU}} A_1(\bar{s},p)\,d\bar{s}\right] \quad (38)$$

where $A_1(s,p)\,ds$ is imaginary at $s = \gamma$. This is because $A_1(\gamma,p)$ is real and $ds = -ids_i$ is imaginary over all of Br_{sL}, hence at $s = \gamma$, since $s = \gamma - is_i$ there, s_i being the imaginary part of s. It follows from the principle of reflection that $A_1(\bar{s},p)\,d\bar{s} = -\overline{A_1(s,p)\,ds}$ and therefore (38) becomes

$$\frac{1}{2\pi i}\int_{Br_{sL}} [A_1(s,p)\,ds - \overline{A_1(s,p)\,ds}] = \frac{1}{\pi}\int_{Br_{sL}} Im[A_1(s,p)\,ds]. \quad (39)$$

According to (37) the second case has the form

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\frac{1}{2\pi i}\left[\int_{Br_{sL}} iA_2(s,p)\,ds+\int_{Br_{sU}} iA_2(\bar{s},p)\,d\bar{s}\right] \tag{40}$$

where $iA_2(s,p)\,ds$ is real at $s=\gamma$, since $A_2(\gamma,p)$ and now ids are real. It follows from the principle of reflection that $iA_2(\bar{s},p)\,d\bar{s}=\overline{iA_2(s,p)\,ds}$ and hence (40) becomes

$$\frac{1}{2\pi i}\int_{Br_{sL}}[iA_2(s,p)\,ds-\overline{iA_2(s,p)\,ds}]$$

$$=\frac{1}{\pi}\int_{Br_{sL}} Im[iA_2(s,p)\,ds]. \tag{41}$$

Summing the right-hand sides of (39) and (41) we have, using the first of (37),

$$\frac{1}{\pi}\int_{Br_{sL}} Im\{[A_1(s,p)+iA_2(s,p)]ds\}=\frac{1}{\pi}\int_{Br_{sL}} Im[A(s,p)\,ds]. \tag{42}$$

Inspection of all the other pairs of terms in (35*b*), and those in (35*a*) (the μ-functions), also conjugate to give (42), including the slightly different pairs of terms in (35*b*) involving $\hat{u}$ and $\hat{v}$. In the case of the latter pairs of terms (42) results through a direct conjugation of the pairs without appeal to the principle of reflection. It follows, therefore, that use of (42) reduces (35*a*) to

$$\begin{Bmatrix}\bar{u}(x,y,p)\\ \bar{v}(x,y,p)\end{Bmatrix}=\frac{1}{\pi}\int_{Br_{sL}} Im\left\{\begin{bmatrix}A_u(s,p)\\ A_v(s,p)\end{bmatrix}ds\right\} \tag{43}$$

where A_u and A_v are the integrands of the first integral on the right-hand side of (35*a*) for $\bar{u}$ and $\bar{v}$, respectively.

Inversion of the Time-Transformed Displacements $\bar{u}(x,y,p)$, $\bar{v}(x,y,p)$ by the Cagniard-deHoop Technique. The formal solutions (43) for $\bar{u}(x,y,p)$, $\bar{v}(x,y,p)$ can be inverted by the Cagniard-deHoop technique (cf. [11], pp. 298–328) since all of the terms in $A_u(s,p)$ and $A_v(s,p)$ in (43) exhibit exponential terms in which the transform parameter p occurs homogeneously linear. The integrands of the first integral on the right-hand side of (35*b*) will show this, and hence (43) will, once the Cagniard-deHoop transformation is invoked.

The Cagniard-deHoop method of inverting (43) begins by introducing the real variable $\zeta=s/k_d$ into the integrals in (43), i.e., with

$$s=k_d\zeta,\quad ds=k_d\,d\zeta,$$

$$\alpha=ik_d(\zeta^2-1)^{1/2},\quad \beta=ik_d(\zeta^2-k^2)^{1/2}, \tag{44}$$

these integrals become

$$\begin{Bmatrix}\bar{u}(x,y,p)\\ \bar{v}(x,y,p)\end{Bmatrix}=\frac{1}{\pi c_d p^2}\int_{Br_{sL}} Im\left\{\left[\begin{Bmatrix}[1\\ -\alpha'/\zeta\end{Bmatrix}\right.\right.$$

$$\left\{[\hat{U}_{dR}(\zeta)-\hat{\mu}_\alpha(\zeta)]e^{-pg_d(\zeta)}+[\pm][\hat{\mu}_{-\alpha}(\zeta)-\hat{\mu}^\infty_{-\alpha}(\zeta)]e^{ph_d(\zeta)}\right\}$$

$$\left.\left.+\begin{Bmatrix}[1\\ \zeta/\beta'\end{Bmatrix}\{[\hat{U}_{sR}(\zeta)-\hat{\mu}_\beta(\zeta)]e^{-pg_s(\zeta)}+[\pm][\hat{\mu}_{-\beta}(\zeta)-\hat{\mu}^\infty_{-\beta}(\zeta)]e^{ph_s(\zeta)}\right\}\right]d\zeta\right\}, \tag{45}$$

where

$$g_d(\zeta)=(1/c_d)[\alpha' y-\zeta x],\quad g_s(\zeta)=(1/c_d)[\beta' y-\zeta x],$$

$$h_d(\zeta)=(1/c_d)[\alpha' y+\zeta x],\quad h_s(\zeta)=(1/c_d)[\beta' y+\zeta x],$$

$$\hat{U}_{dR}(\zeta)=p^3U_{dR}(\zeta)=\{1/R(\zeta)\}\{[(k^2-2\zeta^2)^2$$
$$-4\zeta^2\alpha'\beta']\hat{\mu}^\infty_{-\alpha}(\zeta)-4\zeta^2(k^2-2\zeta^2)\hat{\mu}^\infty_{-\beta}(\zeta)$$
$$+c_d\zeta[(k^2-2)(k^2-2\zeta^2)\hat{u}-2\zeta\beta'\hat{v}]\},$$

$$\hat{U}_{sR}(\zeta)=p^3U_{sR}(\zeta)=-\{1/R(\zeta)\}\{4(k^2-2\zeta^2)\alpha'\beta'\hat{\mu}^\infty_{-\alpha}(\zeta)$$
$$+[(k^2-2\zeta^2)^2-4\zeta^2\alpha'\beta']\hat{\mu}^\infty_{-\beta}(\zeta)$$
$$+c_d\beta'[2(k^2-2)\zeta\alpha'\hat{u}+(k^2-2\zeta^2)\hat{v}]\},$$

$$\hat{\mu}_{\pm\alpha}(\zeta)=p^3\mu_{\pm\alpha}(\zeta)=(\zeta/2k^2)\Big\{\{[(k^2\zeta^2/\alpha')\pm(k^2-2)]A$$
$$-\zeta[(1/\alpha')\pm1]B\}[\pi c_d/(1\mp\alpha')]^{1/2}$$
$$\mp\{c_d[(k^2\zeta^2/\alpha')\pm(k^2-2)c_dc^{-1}_{R1}]C$$
$$-[(c_d/c_{R1}\alpha')\pm1]D\}\{1-\exp[-k_{R1}\mp k_d\alpha')y]\}/$$
$$[\alpha'\mp c_dc^{-1}_{R1}]\pm[c_d^2\sigma_0\zeta/\mu(\zeta^2-1)][1-\exp(\pm k_d\alpha' y)]\Big\},$$

$$\hat{\mu}_{\pm\beta}(\zeta)=p^3\mu_{\pm\beta}(\zeta)=(1/2k^2)\Big\{\{\zeta[k^2\beta'\mp(k^2-2)]A$$
$$-[\beta'\mp\zeta^2]B\}[\pi c_d/(1\mp\beta')]^{1/2}$$
$$\mp c_d\{\zeta[k^2\beta'\mp(k^2-2)c_dc^{-1}_{R1}]C$$
$$-[c_dc^{-1}_{R1}\beta'\mp\zeta^2]D\}\{1-\exp[-(k_{R1}\mp k_d\beta')y]\}/[\beta'\mp c_dc^{-1}_{R1}]$$
$$\mp[c_d^2\sigma_0/\mu][1-\exp(\pm k_d\beta' y)]\Big\},$$

and

$$\hat{\mu}^\infty_{-\alpha}(\zeta)=p^3\mu^\infty_{-\alpha}(\zeta)=(\zeta/2k^2)\Big\{\{[(k^2\zeta^2/\alpha')$$
$$-(k^2-2)]A-\zeta[(1/\alpha')-1]B\}[\pi c_d/(1+\alpha')]^{1/2}$$
$$+c_d\{[(k^2\zeta^2/\alpha')-(k^2-2)c_dc^{-1}_{R1}]C-\zeta[(c_d/c_{R1}\alpha')-1]D\}$$
$$/[\alpha'+c_dc^{-1}_{R1}]-c_d^2\sigma_0\zeta/\mu(\zeta^2-1)\Big\},$$

$$\hat{\mu}^\infty_{-\beta}(\zeta)=p^3\mu^\infty_{-\beta}(\zeta)=(1/2k^2)\Big\{\{\zeta[k^2\beta'+(k^2-2)]A$$
$$-[\beta'+\zeta^2]B\}[\pi c_d/(1+\beta')]^{1/2}$$
$$+c_d\{\zeta[k^2\beta'+(k^2-2)c_dc^{-1}_{R1}]C-[c_dc^{-1}_{R1}\beta'+\zeta^2]D\}$$
$$/[\beta'+c_dc^{-1}_{R1}]+c_d^2\sigma_0/\mu\Big\},$$

$$R(\zeta)=(k^2-2\zeta^2)^2+4\zeta^2\alpha'\beta',\alpha'(\zeta)=\alpha(\zeta)/k_d=i(\zeta^2-1)^{1/2}$$

$$\beta'(\zeta)=\beta(\zeta)/k_d=i(\zeta^2-k^2)^{1/2}.$$

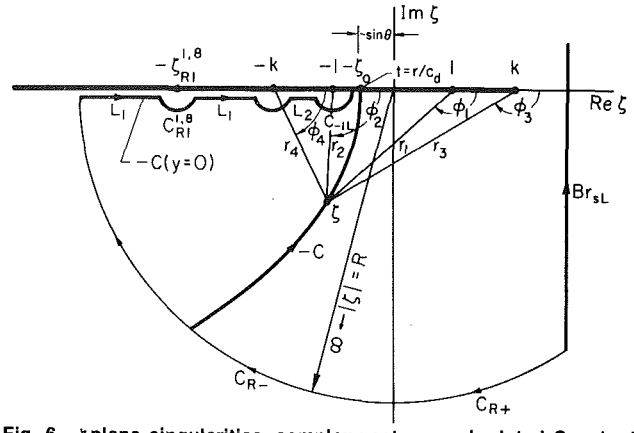


Fig. 6 ζ-plane singularities, complex vectors, and related Cagniard-deHoop integration contours

Inversion of the integrals in (45) by the Cagnaird-deHoop method proceeds by introducing transformations of the type $g(\zeta)=t$ where t is real and positive. These transformations deform the paths of integrations off the path Br_{sL}. Figure 6, stemming from Fig. 5, depicts the ζ-plane with its singularities, branch cuts, contours, and complex vectors. Branch points are at $\zeta=\pm1$ and $\zeta=\pm k$, with the four related branch cuts being on the real axis from k, 1, -1, and $-k$ to $-\infty$, respectively. The Rayleigh function zeros at $-\zeta^{I}_{R1}$ and $-\zeta^{\theta}_{R1}$ generate two simple poles there. This was pointed out earlier, in the discussion of Fig. 4. These zeros correspond to the $-k_{R1}$ pairs shown in Fig. 3 which are associated with

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

paths L_1 and l_{10} and L_8 and l_5, respectively. The superscripts 1 and 8 on $-\zeta_{R1}$ in Fig. 6 indicate the two contributing paths L_1 and L_8 in the present conjugated form of solutions (45). Later we will show these two contributions are equal, and that they add to give the Rayleigh surface wavefronts.

Now from (7) and (8), we have

$$\alpha'(\zeta) = i(\zeta^2 - 1)^{1/2} = (r_1 r_2)^{1/2} e^{i(\psi + \pi/2)}, \quad \psi = (\varphi_1 + \varphi_2)/2,$$

$$\beta'(\zeta) = i(\zeta^2 - k^2)^{1/2} = (r_3 r_4)^{1/2} e^{i(\eta + \pi/2)}, \quad \eta = (\varphi_3 + \varphi_4)/2, \quad (46)$$

where the $r_i = \rho_i / k_d$ $(i = 1, 2, 3, 4)$ are defined in (7) and (8). From (46) it is not difficult to show that

$$Re\left[\frac{\alpha'(\zeta)}{\beta'(\zeta)}\right] > 0, \quad Im\left[\frac{\alpha'(\zeta)}{\beta'(\zeta)}\right] \leq 0 \quad (47)$$

in the third quadrant of the ζ-plane (cf. Fig. 6), and $Re\ \zeta$ and $Im\ \zeta$ are ≤ 0 there. These facts establish the convergence of the integrals associated with $\exp[-pg_d(\zeta)]$ and $\exp[-pg_s(\zeta)]$ in (45), written for paths in the third quadrant. First, for example, we have to show that the integrals for $\bar{u}(x,y,p)$, involving these exponentials over the paths C_{R-} and C_{R+} in Fig. 6, vanish ($\bar{v}(x,y,p)$ is only trivially different). We have from (45)

$$\pi c_d p^2 \bar{u}(x,y,p)\,|_{C_{R-}}$$

$$\lim_{R\to\infty} \int_{C_{R-}} Im\left\{[F_d(\zeta) e^{-pg_d(\zeta)} + F_s(\zeta) e^{-pg_s(\zeta)}] d\zeta\right\} \quad (48)$$

where $F_d(\zeta) = \hat{U}_{dR}(\zeta) - \hat{\mu}_\alpha(\zeta)$, $F_s(\zeta) = \hat{U}_{sR}(\zeta) - \hat{\mu}_\beta(\zeta)$, and $R = |\zeta|$. The contribution of the integral in (48) is assessed by employing Jordan inequalities in the usual way. Hence, from (48)

$$\left|\lim_{R\to\infty} \int_{C_{R-}} Im\left\{[F_d(\zeta) e^{-pg_d(\zeta)} + F_s(\zeta) e^{-pg_s(\zeta)}] d\zeta\right\}\right|$$

$$= \lim_{R\to\infty} \left| \int_{-\pi/2}^{-\pi} Im\left(\left\{ F_d(\zeta) e^{-k_d[\alpha'(\zeta) y - \zeta x]} + F_s(\zeta) e^{-k_d[\beta'(\zeta) y - \zeta x]} \right\}_{\zeta = Re^{i\varphi}} ie^{i\varphi} \right) R d\varphi \right|$$

$$\leq \lim_{R\to\infty} \left[R \int_0^{\pi/2} \left| Im\left\{ F_d(Re^{i\varphi}) + F_s(Re^{i\varphi}) \right] ie^{i[k_d R(x\sin\varphi - y\cos\varphi) + \varphi]} \right\} \right| e^{k_d R(x\cos\varphi + y\sin\varphi)} d\varphi \Bigg]. \quad (49)$$

The last integral in (49) reduces to

$$\lim_{R\to\infty} \left[\frac{K}{R^{\kappa-1}} e^{-k_d x R} \int_0^{\pi/2} e^{2k_d(x-y)R\varphi/\pi} d\varphi \right]$$

$$= \lim_{R\to\infty} \left[\frac{\pi K}{2k_d(x-y)R^\kappa} (e^{-k_d y R} - e^{-k_d x R}) \right] \to 0 \quad (50)$$

where κ and K (both real) define the order of, and constant associated with, the component terms of $F_d(Re^{i\varphi})$ $+F_s(Re^{i\varphi})$, respectively. The order index κ satisfies the requirement $\kappa > 0$ since it takes on the values 1/2, 1, 3/2, and 2. It follows from (50), therefore, that the integral in (48) also vanishes. Note for $x > 0, y > 0$ (50) holds without restriction. For $x > 0$, $y = 0$ or $y > 0$, $x = 0$ (50) holds, however it reduces to the usual one-dimensional form in each case. Finally, when $x = y = 0$ (the vertex of the quarter plane) (50) becomes indeterminate and does not hold. However, $\bar{u}(0,0,p)$ and $\bar{v}(0,0,p)$ are given by (33) and $\hat{u}$ and $\hat{v}$ there are defined by $(A.2)$ and $(A.1)$ of the Appendix.

Similarly, we can show the counterpart integral of (48) over the path C_{R+} also vanishes. Except for a difference in integration limits (here, between $-\cos^{-1}(\gamma/R)$ and $-\pi/2$) we again have the equation (49) but now for the path C_{R+}. Through the transformation $q = \cos\varphi$, the last integral in (49) for the present case becomes

$$\lim_{R\to\infty} \left[\frac{K}{R^{\kappa-1}} \int_0^{\gamma/R} e^{k_d R[xq - y(1-q^2)^{1/2}]} \frac{dq}{(1-q^2)^{1/2}} \right],$$

and since $q << 1$ when $R \to \infty$ this integral reduces to the approximation

$$\lim_{R\to\infty} \left[\frac{K}{R^{\kappa-1}} e^{-k_d y R} \int_0^{\gamma/R} e^{k_d x R q} dq \right]$$

$$= \lim_{R\to\infty} \left[\frac{K}{k_d x R^\kappa} e^{-k_d y R} (e^{k_d x \gamma} - 1) \right]$$

$$= \lim_{R\to\infty} \left[\frac{K}{k_d x R^\kappa} e^{-k_d y R} (e^{k_d x \gamma} - 1) \right] \to 0, \quad (51)$$

and it follows from (51) that the integral in (48) for the C_{R+} path also vanishes. The limit (51) holds for all cases, i.e., $x > 0, y > 0$; $x > 0, y = 0$; $y > 0, x = 0$; $x = y = 0$.

The contributions of the integrals in (45), associated with $\exp[ph_d(\zeta)]$ and $\exp[ph_s(\zeta)]$ and over the paths C_{R-} and C_{R+}, also vanish. The analysis involving equations (48)–(51) can be used to show this. First we note that $F_d(\zeta) e^{ph_d(\zeta)} + F_s(\zeta) e^{ph_s(\zeta)}$ in (48), where $F_d(\zeta) = \hat{\mu}_{-\alpha}(\zeta) - \hat{\mu}^\infty_{-\alpha}(\zeta)$ and $F_s(\zeta) = \hat{\mu}_{-\beta}(\zeta) - \hat{\mu}^\infty_{-\beta}(\zeta)$ are defined through (45), reduces to

$$-\frac{1}{2k^2} \left\{ \frac{\zeta\{c_d[(k^2\zeta^2/\alpha') - (k^2-2)c_d c_{R1}^{-1}]C - \zeta[(c_d/c_{R1}\alpha') - 1]D\} e^{-k_{R1} y}}{\alpha' + c_d c_{R1}^{-1}} \right.$$

$$+ \frac{c_d\{\zeta[k^2\beta' + (k^2-2)c_d c_{R1}^{-1}]C - [c_d c_{R1}^{-1}\beta' + \zeta^2]D\} e^{-k_{R1} y}}{\beta' + c_d c_{R1}^{-1}}$$

$$\left. - \frac{c_d^2 \sigma_0 \zeta^2}{\mu(\zeta^2 - 1)} + \frac{c_d^2 \sigma_0 \zeta^2}{\mu} \right\} e^{k_d x \zeta}.$$

It follows, using this expression for the absolute value of the integral over C_{R-} that (50) governs, gives the result

$$\lim_{R\to\infty} \left[\frac{\pi K e^{-k_{R1} y}}{2k_d x R^\kappa} (1 - e^{-k_d x R}) \right] \to 0 \quad (52)$$

where $\kappa = 1$ and K contains the multiplier C. Because of the result (52) the integral over C_{R-} also vanishes.

It also follows for the absolute value of the integral over C_{R+} that (51) holds in the present case, giving the result

$$\lim_{R\to\infty} \left[\frac{K e^{-k_{R1} y}}{k_d x R^\kappa} (e^{k_d x \gamma} - 1) \right] \to 0 \quad (53)$$

where κ and K are as stated in (52). The result (53) proves that the integral over C_{R+} also vanishes.

In Part 2 of this paper all of the quarter-plane wavefront events in the edge uniform pressure problem (cf. Fig. 1) are

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

derived through further steps in the Cagniard-deHoop inversion of (45) and related asymptotics.

References

1 Knopoff, L., "Elastic Wave Propagation in a Wedge," *Wave Propagation in Solids*, J. Miklowitz, Ed., ASME Publications, New York, 1969, pp. 3-43.

2 Alterman, Z. S., and Rotenberg, A., "Seismic Waves in a Quarter Plane," *Bull. Seism. Soc. Amer.*, Vol. 59, 1969, pp. 347-368.

3 Achenbach, J. D., and Khetan, R. P., "Elastodynamic Response of a Wedge to Surface Pressures," *International Journal of Solids and Structures*, Vol. 13, 1977, pp. 1157-1171.

4 Achenbach, J. D., "Wave Propagation, Elastodynamic Stress Singularities, and Fracture," *Theoretical and Applied Mechanics*, Koiter, W. T., Ed., North-Holland, Amsterdam, 1976, pp. 71-88.

5 Wojcik, G. L., "Self-Similar Elastodynamic Solutions for the Plane Wedge," Ph.D. Thesis, California Institute of Technology, Pasadena, Calif., 1977. Also available as a Graduate Aeronautical Laboratories Report, C.I.T.

6 Lapwood, E. R., "The Transmission of a Rayleigh Pulse Round a Corner," *Geophys. Journal*, Vol. 4, 1961, pp. 174-196.

7 Miklowitz, J., "Elastic Wave Propagation," *Applied Mechanics Surveys*, Spartan Books, Washington, D.C., 1966, p. 827.

8 Miklowitz, J., "Analysis of Elastic Waveguides Involving an Edge," *Wave Propagation in Solids*, ASME Publications, New York, 1969, pp. 44-70.

9 Sinclair, G. B., and Miklowitz, J., "Two Nonmixed Symmetric End Loadings of an Elastic Waveguide," *Int. Journal Solids Structures*, Vol. 11, 1975, pp. 275-294.

10 Miklowitz J., and Garrott, W. R., "Transient Response of Two-Dimensional Cantilevered Plates Subjected to Base Motions," *Proceedings IUTAM Symposium on Modern Problems in Elastic Wave Propagation*, Miklowitz J., and Achenbach, J. D., Eds., Wiley-Interscience, New York, 1978.

11 Miklowitz, J., "*The Theory of Elastic Waves and Waveguides*," North Holland, Amsterdam, New York, and Oxford, 1978.

12 Rosenfeld, R. L., and Miklowitz, J., "Wavefronts in Elastic Rods and Plates," *Proc. Fourth U.S. Nat. Congr. Appl. Mech.*, Vol. 1, ASME, New York, 1962, pp. 293-303.

APPENDIX

Definitions of the constants A, B, C, and D in the wavefront p-dependent coefficients given by (29):

$$A=[(\hat{c}_2-\hat{c}_3)/W][-\hat{T}_1(\hat{B}_2\hat{D}_3-\hat{B}_3\hat{D}_2)+\hat{T}_2(b_u\hat{D}_3-d_u\hat{B}_3)+\hat{T}_3(d_u\hat{B}_2-b_u\hat{D}_2)],$$

$$B=W^{-1}\Big\{(\hat{a}_2-\hat{a}_3)[\hat{T}_2\{\hat{D}_3(\hat{c}_3+c_u)-\hat{S}_3c_vd_u\}-\hat{S}_2c_v\{\hat{T}_1\hat{D}_3-\hat{T}_3d_u\}+D_2\{\hat{T}_1\hat{S}_3c_v-\hat{T}_3(\hat{c}_3+c_u)\}]+(\hat{c}_2-\hat{c}_3)[\hat{S}_2a_v\{\hat{T}_1\hat{D}_3-\hat{T}_3d_u\}+\hat{T}_2\{\hat{D}_3(\hat{a}_3+a_u)-\hat{S}_3a_vd_u\}+\hat{D}_2\{\hat{T}_3(\hat{a}_3+a_u)-\hat{T}_1\hat{S}_3a_v\}]\Big\},$$

$$C=-[(\hat{a}_2-\hat{a}_3)/(\hat{c}_2-\hat{c}_3)]A,$$

and

$$D=-W^{-1}\Big\{(\hat{a}_2-\hat{a}_3)[\hat{T}_2\{\hat{B}_3(\hat{c}_3+c_u)-\hat{S}_3c_vb_u\}-\hat{S}_2c_v\{\hat{T}_1\hat{B}_3-\hat{T}_3b_u\}+\hat{B}_2\{\hat{T}_1\hat{S}_3c_v-\hat{T}_3(\hat{c}_3+c_u)\}]+(\hat{c}_2-\hat{c}_3)[\hat{S}_2a_v\{\hat{T}_1\hat{B}_3-\hat{T}_3b_u\}+\hat{T}_2\{\hat{B}_3(\hat{a}_3+a_u)-\hat{S}_3a_vb_u\}+\hat{B}_2\{\hat{T}_3(\hat{a}_3+a_u)-\hat{T}_1\hat{S}_3a_v\}]\Big\}, \quad (A.1)$$

where

$$W=\Big\{(\hat{a}_2-\hat{a}_3)[\hat{B}_2\{\hat{D}_3(\hat{c}_3+c_u)-\hat{S}_3c_vd_u\}-b_uc_v\{\hat{D}_3\hat{S}_2-\hat{D}_2\hat{S}_3\}+\hat{B}_3\{\hat{S}_2c_vd_u-\hat{D}_2(\hat{c}_3+c_u)\}]+(\hat{c}_2-\hat{c}_3)[\hat{S}_2a_v\{b_u\hat{D}_3-d_u\hat{B}_3\}-(\hat{a}_3+a_u)\{\hat{B}_2\hat{D}_3-\hat{B}_3\hat{D}_2\}+\hat{S}_3a_v\{d_u\hat{B}_2-b_u\hat{D}_2\}]\Big\},$$

with, for $j=2,3$,

$$\hat{B}_j=\hat{b}_j+\hat{S}_jb_v,\quad \hat{D}_j=\hat{d}_j+\hat{S}_jd_v,$$

$$\hat{T}_j=(2-s_{Rj}a_j^{-2})\sigma_0/\mu c_{Rj};\quad \hat{T}_1=-(2+s_{R1}a_1^{-2})\sigma_0/\mu c_{R1},$$

and

$$\hat{a}_j=\pi^{1/2}c_{Rj}^{-2}[s_{Rj}f_j^{-1}\{k^2a_j^{-1}-c_d^{-1}(k^2-2)c_{Rj}^2\}+2g_j^{-1}\{k^2b_j-c_d^{-1}(k^2-2)\}],$$

$$\hat{b}_j=\pi^{1/2}c_{Rj}^{-1}[s_{Rj}f_j^{-1}\{1-(c_da_j)^{-1}\}-2g_j^{-1}\{c_d^{-1}b_j-c_{Rj}^{-2}\}],$$

$$\hat{c}_j=c_{Rj}^{-2}[s_{Rj}h_j^{-1}\{k^2a_j^{-1}-c_{R1}^{-1}(k^2-2)c_{Rj}^2\}+2i_j^{-1}\{k^2b_j-c_{R1}^{-1}(k^2-2)\}],$$

$$\hat{d}_j=c_{Rj}^{-1}[s_{Rj}h_j^{-1}\{1-(c_{R1}a_j)^{-1}\}-2i_j^{-1}\{c_{R1}^{-1}b_j-c_{Rj}^{-2}\}],$$

$$\hat{S}_j=-a_1s_{Rj}c_{Rj}/s_{R1}c_{R1}a_j,$$

with

$$a_j=(c_d^{-2}-c_{Rj}^{-2})^{1/2};\quad a_1=(c_{R1}^{-2}-c_d^{-2})^{1/2},$$

$$b_j=(c_s^{-2}-c_{Rj}^{-2})^{1/2};\quad b_1=(c_{R1}^{-2}-c_s^{-2})^{1/2},$$

$$s_{Rj}=c_s^{-2}-2c_{Rj}^{-2};\quad s_{R1}=c_s^{-2}-2c_{R1}^{-2},\quad c_{Rj}=\gamma_jc_s;\quad c_{R1}=\gamma_1c_s,$$

$$f_j=(a_j+c_d^{-1})^{1/2},\quad g_j=(b_j+c_d^{-1})^{1/2},$$

$$h_j=a_j+c_{R1}^{-2},\quad i_j=b_j+c_{R1}^{-2},$$

and where

$$a_u=(\pi/2)^{1/2}c_{R1}^{-2}[c_{R1}s_{R1}\{k^2c_2a_1^{-1}+(k^2-2)c_{R1}^2c_d^{-1}c_1\}+\{k^2c_{R1}^2b_1d_2+2(k^2-2)(c_dq)^{-1}d_1\}],$$

$$a_v=(\pi/2)^{1/2}c_{R1}^{-2}[c_{R1}s_{R1}\{k^2c_1a_1^{-1}-(k^2-2)c_{R1}^2c_d^{-1}c_2\}+2q^{-1}\{k^2b_1d_1-(k^2-2)c_d^{-1}d_2\}],$$

$$b_u=-(\pi/2)^{1/2}[s_{R1}\{(c_da_1)^{-1}c_2+c_1\}+2(c_{R1}q)^{-1}\{c_d^{-1}b_1d_2+c_{R1}^{-2}d_1\}],$$

$$b_v=-(\pi/2)^{1/2}[s_{R1}\{(c_da_1)^{-1}c_1-c_2\}+2(c_{R1}q)^{-1}\{c_d^{-1}b_1d_1-c_{R1}^{-2}d_2\}],$$

$$c_u=-2c_{R1}^{-2}[(k^2-1)r_{R1}^{-1}s_{R1}+s_{R1}^{-1}\{k^2b_1^2+(k^2-2)c_{R1}^{-2}\}],$$

$$c_v=-c_{R1}^{-3}[r_{R1}^{-1}s_{R1}\{k^2a_1^{-1}-(k^2-2)c_{R1}^2a_1\}+4s_{R1}^{-1}b_1],$$

$$d_u=2c_{R1}^{-2}r_{R1}^{-1}(s_{R1}-r_{R1}),\quad d_v=c_d^{-2}(c_{R1}a_1)^{-1}r_{R1}^{-1}s_{R1},$$

with

$$c_l=[c_{R1}^{-1}+(-)^{l-1}c_d^{-1}]^{1/2},\quad d_l=[q+(-)^{l-1}c_d^{-1}]^{1/2},$$

$l=1,2$,

$$r_{R1}=c_d^{-2}-2c_{R1}^{-2}\quad\text{and}\quad q=(b_1^2+c_d^{-2})^{1/2},$$

and c_d, c_s, and k are defined after (1).

Definitions of the constants $\hat{u}$ and $\hat{v}$ in the p-dependent corner displacements given by (33):

$$\hat{u}=c_s^2(k^2-2)^{-1}[-\hat{T}_1+a_uA+b_uB+c_uC+d_uD],$$

and

$$\hat{v}=-2c_s^2a_1(c_{R1}s_{R1})^{-1}[a_vA+b_vB+c_vC+d_vD], \quad (A.2)$$

where a_u, a_v, b_u, b_v, c_u, c_v, d_u, d_v, A, B, C, D, $\hat{T}_1$, a_1, c_{R1}, and s_{R1} are given in (A.1), and c_s and k are defined after (1).

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

J. Miklowitz
Professor,
Division of Engineering and
Applied Science,
California Institute of Technology,
Pasadena, Calif. 91125
Fellow ASME

Wavefront Analysis in the Nonseparable Elastodynamic Quarter-Plane Problems, I

Part 2: Wavefront Events in the Edge Uniform Pressure Problem

Introduction

As pointed out in Part 1 of this paper, here in Part 2 all of the quarter-plane wavefront events in the edge uniform pressure problem (cf. Fig. 1, Part 1) are derived through further steps in the Cagniard-deHoop inversion of (45, Part 1) and a related asymptotic method due to Rosenfeld and Miklowitz [12]. In the interior of the quarter plane the events include, for the displacements, velocities, and accelerations, the regular dilatational wavefronts, the two-sided equivoluminal and head wavefronts in the two critical regions associated with the edges (surfaces) $x = 0$ and $y = 0$, and of course the plane dilatational wavefront from the edge input. The events on the edges of the quarter plane include, for $y = 0$, the wavefronts just discussed for the interior and, in addition, the Rayleigh wavefronts. Concerning the loaded edge $x = 0$, (30, Part 1) and (31, Part 1) have already defined the dilatational and Rayleigh wavefronts. However, as pointed out in Part 1, the present asymptotic method also yields the two-sided equivoluminal and head wavefronts on the loaded edge, $x = 0$.

Wavefront Events for the Displacements, Velocities, and Accelerations in the Quarter-Plane Interior

The Dilatational Wavefronts. Let $g_d(\zeta) = t$(real), which from (45, Part 1) gives

$$t = g_d(\zeta) = (1/c_d)[i(\zeta^2 - 1)^{1/2} y - \zeta x]. \tag{1}$$

Transformation (1) deforms path Br_{SL} into the equivalent path $-C$ (C for Cagniard-deHoop), shown in Fig. 6, Part 1, which is defined by solving (1) for $\zeta = \zeta_d(t)$, where

$$\zeta_d(t) = -[(c_d t/r)\sin\theta + i\{(c_d t/r)^2 - 1\}^{1/2}\cos\theta],$$
$$r/c_d \le t < \infty, \quad 0 \le \theta \le \pi/2 \tag{2a}$$

in which $x = r\sin\theta$, $y = r\cos\theta$, and $r = (x^2 + y^2)^{1/2}$. It is easy to show that (2a) representing path C, is one-half of the branch of the hyperbola

$$\left(\frac{Re\,\zeta_d}{\sin\theta}\right)^2 - \left(\frac{Im\,\zeta_d}{\cos\theta}\right)^2 = 1 \tag{2b}$$

in the third quadrant of the ζ-plane. The earliest time in (2a) is $t = r/c_d$ (the wavefront arrival time), which when substituted into (2a) gives $\zeta_d = -\zeta_0 = -\sin\theta$, the point of intersection of the beginning of path C and the negative ζ-axis as Fig. 6, Part 1 indicates. Equation (2a) shows that as time grows from $t = r/c_d$, ζ_d moves on C outward into the third quadrant, approaching for large time the asymptote defined by $\lim(Re\,\zeta_d/Im\,\zeta_d) = \tan\theta$. The dilatational wavefronts are extracted from the first integrals in (45, Part 1) on path $-C$,

$$\begin{bmatrix} \bar{u}_d(x,y,p) \\ \bar{v}_d(x,y,p) \end{bmatrix} = \frac{1}{\pi c_d p^2}\int_{-C}^{\infty} Im\left\{\begin{Bmatrix} 1 \\ -\alpha'/\zeta \end{Bmatrix}\{[\hat{U}_{dR}(\zeta) - \hat{\mu}_\alpha(\zeta)]e^{-pg_d(\zeta)}\}d\zeta\right\}. \tag{3}$$

Now substituting $g_d(\zeta) = t$ from (1) and $\zeta_d(t)$ from (2a) into (3) transforms the latter into

$$\begin{bmatrix} \bar{u}_d(x,y,p) \\ \bar{v}_d(x,y,p) \end{bmatrix} = -\frac{1}{\pi c_d p^2}\int_{r/c_d}^{\infty} Im\left\{\begin{Bmatrix} 1 \\ -\alpha'(\zeta_d)/\zeta_d \end{Bmatrix}\times\{[\hat{U}_{dR}(\zeta_d) - \hat{\mu}_\alpha(\zeta_d)]/g_d'(\zeta_d)\}\right\}e^{-pt}\,dt \tag{4}$$

where we have used $d\zeta_d(t)/dt = 1/g_d'[\zeta_d(t)]$, the prime on g_d indicating first derivative, and we have dropped the argument t of ζ_d. By inspection of (4) we note it yields the dilatational acceleration components

$$\begin{bmatrix} \ddot{u}_d(x,y,t) \\ \ddot{v}_d(x,y,t) \end{bmatrix} = -\frac{1}{\pi c_d} Im\left\{\begin{Bmatrix} 1 \\ -\alpha'(\zeta_d)/\zeta_d \end{Bmatrix}\{[\hat{U}_{dR}(\zeta_d) - \hat{\mu}_\alpha(\zeta_d)]/g_d'(\zeta_d)\}\right\}\times H(t - r/c_d), \tag{5}$$

H being the Heaviside step function.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1981; final revision, April, 1982. Paper No. 82-WA/APM-13.
Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

To approximate the dilatational acceleration components in (5) near their fronts, we note that there $t = g_d(\zeta_d) = r/c_d + \epsilon$ where $0 \le \epsilon << 1$. The point $\zeta_d = -\zeta_0 = -\sin\theta$ satisfies $g_d'(-\zeta_0) = 0$, and hence $\zeta_d = -\zeta_0$, corresponding to the minimum time r/c_d, is a saddle point of the relief of $\exp[-pg_d(\zeta_d)]$ in (3), where now p is large. The method of steepest descents could be employed at this point, but the technique in [12] is much simpler. The latter begins with a Taylor series expansion of $g_d(\zeta_d)$ at $\zeta_d = -\zeta_0$ along the path C with the result

$$t = g_d(\zeta_d) = g_d(-\zeta_0) + [g_d''(-\zeta_0)/2](\zeta_d+\zeta_0)^2 + \cdots$$
$$= r/c_d + [g_d''(-\zeta_0)/2](\zeta_d+\zeta_0)^2 + \cdots . \quad (6)$$

From (6)

$$\zeta_d + \zeta_0 \simeq [2(t-r/c_d)/g_d''(-\zeta_0)]^{1/2}$$

and

$$g_d'(\zeta_d) \simeq g_d''(-\zeta_0)(\zeta_d+\zeta_0)$$

from which

$$g_d'(\zeta_d) \simeq [2g_d''(-\zeta_0)(t-r/c_d)]^{1/2} \quad (7)$$

where $g_d''(-\zeta_0) = i^2 r/c_d \cos^2\theta$. It follows from (7) and setting $\hat{U}_{dR}(\zeta_d) - \hat{\mu}_\alpha(\zeta_d) \simeq \hat{U}_{dR}(-\zeta_0) - \hat{\mu}_\alpha(-\zeta_0)$, that (5) is approximated by the real results

$$\begin{bmatrix} \ddot{u}_d(x,y,t) \\ \ddot{v}_d(x,y,t) \end{bmatrix} \simeq \frac{\cos\theta}{\pi(2c_d r)^{1/2}} \begin{bmatrix} 1 \\ \cot\theta \end{bmatrix} [\hat{U}_{dR}(-\zeta_0) - \hat{\mu}_\alpha(-\zeta_0)]$$
$$\times (t-r/c_d)^{-1/2} H(t-r/c_d) \quad (8)$$

where

$$\hat{U}_{dR}(-\zeta_0) = \{[(k^2-2\sin^2\theta)^2 - 4\sin^2\theta\cos\theta(\sin^2\theta-k^2)^{1/2}]\hat{\mu}^{\infty}_{\mu-\alpha}(-\zeta_0)$$
$$+ 4\sin^2\theta(k^2-2\sin^2\theta)\hat{\mu}^{\infty}_{-\beta} - c_d\sin\theta[(k^2-2)(k^2-2\sin^2\theta)$$
$$\times \hat{u} + 2\sin\theta(k^2-\sin^2\theta)^{1/2}\hat{v}]\}/R(-\zeta_0),$$

$$\hat{\mu}_\alpha(-\zeta_0) = -\frac{\sin\theta}{2k^2}\left\{\left[\left(\frac{k^2\sin^2\theta}{\cos\theta} + k^2 - 2\right)A + \sin\theta\left(\frac{1}{\cos\theta}+1\right)B\right]\left(\frac{c_d\pi}{1-\cos\theta}\right)^{1/2}\right\},$$

and

$$\hat{\mu}^{\infty}_{-\alpha}(-\zeta_0) = -\frac{\sin\theta}{2k^2}\left\{\left[A\left(\frac{k^2\sin^2\theta}{\cos\theta} - (k^2-2)\right) + B\sin\theta\left(\frac{1}{\cos\theta}-1\right)\right]\left(\frac{\pi c_d}{\cos\theta+1}\right)^{1/2}\right.$$
$$+\left[C\left(\frac{k^2\sin^2\theta}{\cos\theta} - (k^2-2)c_d C_{R1}^{-1}\right) + D\sin\theta\left(\frac{c_d}{C_{R1}\cos\theta}-1\right)\right]$$
$$\left.\times \frac{c_d}{\cos\theta + c_d C_{R1}^{-1}} - \frac{c_d^2\sigma_0\sin\theta}{\mu\cos^2\theta}\right\},$$

$$\hat{\mu}^{\infty}_{-\beta}(-\zeta_0) = \frac{1}{2k^2}\left\{[A\{k^2(k^2-\sin^2\theta)^{1/2} - (k^2-2)\sin\theta\}\right.$$
$$- B\{\sin^2\theta + (k^2-\sin^2\theta)^{1/2}\}]\left(\frac{\pi c_d}{(k^2-\sin^2\theta)^{1/2}+1}\right)^{1/2}$$
$$- [C\sin\theta\{k^2(k^2-\sin^2\theta)^{1/2} + (k^2-2)c_d c_{R1}^{-1}\}$$
$$\left. - D\{c_d c_{R1}^{-1}(k^2-\sin^2\theta)^{1/2}\}]\frac{c_d}{(k^2-\sin^2\theta)^{1/2} + c_d c_{R1}^{-1}} + \frac{c_d^2\sigma_0}{\mu}\right\},$$

$$R(-\zeta_0) = (k^2-2\sin^2\theta)^2 + 4\sin^2\theta\cos\theta(k^2-\sin^2\theta)^{1/2}.$$

The velocity and displacement wavefront components are obtained by simple successive time integrations of (8). The results are (8) except for a replacement of $(t-r/c_d)^{-1/2}$ there, by $2(t-r/c_d)^{1/2}$ for the velocities $\dot{u}_d(x,y,t)$ and $\dot{v}_d(x,y,t)$, and by $(4/3)(t-r/c_d)^{3/2}$ for the displacements $u_d(x,y,t)$ and $v_d(x,y,t)$.

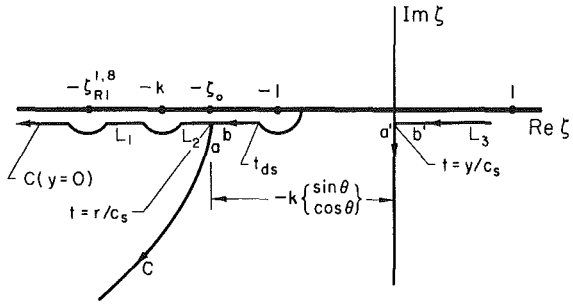
It is important to note that the approximation (8) for $\ddot{u}_d(x,y,t)$ (as well as $\dot{u}_d, u_d$) breaks down at the free edge $y=0$ ($\theta=\pi/2$) through an infinite value caused by the last term in $\hat{\mu}^{\infty}_{-\alpha}(-\zeta_0)$. Also, at the loaded edge $x=0(\theta=0)$, this approximation fails through a vanishing value. On the other hand, approximation (8) for $\ddot{v}_d(x,y,t)$ (as well as $\dot{v}_d, v_d$) does hold on both of the quarter-plane edges.

The Equivoluminal Wavefronts. The process for deriving the present wavefronts is basically the same as that used for the foregoing dilatational wavefronts. Here it is applied to the $g_s(\zeta)$ integrals in (45, Part 1). We now let

$$t = g_s(\zeta) = (1/c_d)[i(\zeta^2-k^2)^{1/2}y - \zeta x]. \quad (9)$$

Transformation (9) deforms path Br_{sL} into the path $-C$ in Fig. 1, which is defined by solving (9) for $\zeta = \zeta_s(t)$, where

$$\zeta_s(t) = -[(c_d t/r)\sin\theta + i\{(c_s t/r)^2 - 1\}^{1/2} k\cos\theta],$$
$$r/c_s \le t < \infty, \quad 0 \le \theta \le \pi/2. \quad (10)$$



θ domain: $I < k\sin\theta < k$ or $\theta_{crx} < \theta < \pi/2$

$I < k\cos\theta < k$ or $0 < \theta < \theta_{cry}$

Fig. 1 Contours for the head and two-sided equivoluminal wavefront approximations

Again (10) is one-half of a branch of a hyperbola [similar to (2*b*)] in the third quadrant of the ζ-plane with the same asymptote as (2*b*). The earliest time in (10) is $t = r/c_s$, which is both the wavefront arrival time for the *regular part* (no precursor involved) of the two-sided equivoluminal wave and the final time associated with the *precursor part* of this wave. Figure 1 shows that $t = r/c_s$, with $\zeta_0 = -k\sin\theta$, the corresponding saddle point, is at the corner of path C. As we shall see, the two-sided wavefront is obtained through expansions along C from $t = r/c_s$, outward from the corner along the hyperbola (a short time, say to $t = r/c_s + \epsilon$) and along the horizontal line (a short time, say to $t = r/c_s - \epsilon$). The first of these expansions gives the regular part, and the second the precursor part of this wavefront event.

As pointed out earlier and shown in Fig. 1, Part 1, equivoluminal waves in the quarter-plane problem occur in two overlapping regions which are associated with the edges (surfaces) $y=0$ (x-axis) and $x=0$ (y-axis). These regions are found from Snell's law for the reflection of *SV* waves at critical angles of incidence, which for the x-axis edge is $\sin\hat{\alpha} = k\sin\hat{\beta}$, $\hat{\alpha}$ and $\hat{\beta}$ being angles of incidence and reflection for *SV* and *P* waves, respectively (cf. §3.1.4.5 in [11]). The critical region is easily seen to be $(1/k) < \sin\hat{\beta} < 1$ which can be written here, replacing $\hat{\beta}$ with θ, as

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\theta_{crx} = \sin^{-1}(1/k) < \theta < \pi/2 \text{ or } 1 < k \sin\theta < k. \quad (11a)$$

Similarly, for the y-axis, the Snell's law is $\cos\hat{\alpha} = k\cos\hat{\beta}$, which leads to the critical region

$$0 < \theta < \theta_{cry} = \cos^{-1}(1/k) \text{ or } 1 < k\cos\theta < k. \quad (11b)$$

Note that the second inequalities in (11a,b) agree with the position of the saddle points in Fig. 1 for the two cases. Note also that the saddle points are on the branch cut from $\zeta = -1$ to $-\infty$, hence the path C has to include a part of path L_2 just below the cut from $-\zeta_0$ to -1 in Fig. 1. This part of path C contributes the head wave, and is typical of this event in critical region analysis for equivoluminal waves. The head wavefronts will be treated first and then the two-sided equivoluminal wavefronts.

The head wavefronts are derived by first noting that on the L_2 part of path C in Fig. 1, ζ becomes real and negative in the limit as the branch cut is approached. Further, $|\zeta|$ decays along this line from $\zeta = -\zeta_0 = -k\sin\theta$ toward its end at the branch point $\zeta = -1$. From (10), then, it follows that

$$\zeta_s(t) = \zeta_{ds}(t) = [1 - (c_s t/r)^2]k\cos\theta - (c_d t/r)\sin\theta \quad (12)$$

showing that $t \le r/c_s$ along this part of path C. The minimum time t_{ds} is reached at the branch point $\zeta = -1$, hence substituting the latter into (9) we find

$$t_{ds} = (1/c_d)[i(1-k^2)^{1/2}y + x]$$
$$= (r/c_d)[(k^2-1)^{1/2}\cos\theta + \sin\theta] \quad (13)$$

since $\beta'(-1) = i(1-k^2)^{1/2} = i[-i(k^2-1)^{1/2}] = (k^2-1)^{1/2}$ in agreement with its positive value on L_2, i.e., $\beta(\zeta) = (\rho_3\rho_4)^{1/2}$.

Now integrals analogous to (3) and (4) are written for the $g_s(\zeta)$ integrals in (45, Part 1), through the transformations (9), (10), and $d\zeta_{ds}(t)/dt = 1/g_s'[\zeta_{ds}(t)]$, which leads to

$$\begin{Bmatrix} \ddot{u}_{ds}(x,y,t) \\ \ddot{v}_{ds}(x,y,t) \end{Bmatrix} = -\frac{1}{\pi c_d} Im\left[\begin{Bmatrix} 1 \\ \zeta_{ds}/\beta'(\zeta_{ds}) \end{Bmatrix} \{[\hat{U}_{sR}(\zeta_{ds}) - \hat{\mu}_\beta(\zeta_{ds})]/g_s'(\zeta_{ds})\} H(t-t_{ds}). \quad (14)$$

To approximate the head equivoluminal acceleration components in (14) near their fronts, we first note that there $t = t_{ds} + \epsilon$ where $0 \le \epsilon << 1$. Now the point ζ_{ds} corresponding to the minimum time t_{ds} is $\zeta_{ds} = -1$, which in fact led to (13). So we expand $g_s(\zeta_{ds})$ at $\zeta_{ds} = -1$ along L_2 (horizontal line part of C) via

$$t = g_s(\zeta_{ds}) = g_s(-1) + g_s'(-1)(\zeta_{ds}+1) + \ldots \quad (15a)$$
$$\simeq t_{ds} + g_s'(-1)(\zeta_{ds}+1), \quad (15b)$$

and from (15b) we have

$$g_s'(-1) = (t - t_{ds})/(\zeta_{ds}+1). \quad (16)$$

Then using (9), and the fact that β' is a positive number on L_2, we find that

$$g_s'(-1) = -(r/c_d)[\sin\theta - (k^2-1)^{-1/2}\cos\theta], \quad (17)$$

a negative quantity in the present region $\theta_{crx} < \theta < \pi/2$. Now near $\zeta_{ds} = -1$, from (16) we have

$$\alpha'(\zeta_{ds}) = i[(\zeta_{ds}-1)(\zeta_{ds}+1)]^{1/2}$$
$$= -i[2(t-t_{ds})/|g_s'(-1)|]^{1/2} \quad (18)$$

in agreement with the value of α' on path L_2, i.e., $\alpha = -i(\rho_1\rho_2)^{1/2}$. It follows from (15$a$) that $g_s'(\zeta_{ds}) \simeq g_s'(-1)$, and substituting this and the approximation $\hat{U}_{sR}(\zeta_{ds}) - \hat{\mu}_\beta(\zeta_{ds}) \simeq \hat{U}_{sR}(-1) - \hat{\mu}_\beta(-1)$ into (14) reduces the latter to

$$\begin{Bmatrix} \ddot{u}_{ds}(x,y,t) \\ \ddot{v}_{ds}(x,y,t) \end{Bmatrix} \simeq -\frac{1}{\pi c_d}\begin{Bmatrix} 1 \\ -(k^2-1)^{-1/2} \end{Bmatrix} Im\{[\hat{U}_{sR}(-1) - \hat{\mu}_\beta(-1)]/g_s'(-1)\} H(t-t_{ds}), \quad (19)$$

where we have used $\zeta_{ds} = -1$ and $\beta' = (k^2-1)^{1/2}$. Now substituting $\alpha'(\zeta_{ds})$, given by the right-hand side of the second equation of (18), and $(k^2-1)^{1/2}$ for β', into $\hat{U}_{sR}(-1) - \hat{\mu}_\beta(-1)$, and making use of (17) for $g'_s(-1)$, (19) reduces to the real results for the acceleration head wavefront components

$$\begin{Bmatrix} \ddot{u}_{ds}(x,y,t) \\ \ddot{v}_{ds}(x,y,t) \end{Bmatrix} = \begin{Bmatrix} 1 \\ -(k^2-1)^{-1/2} \end{Bmatrix} \{f(\theta)[r(t-t_{ds})]^{-1/2} + 0[(t-t_{ds})^{1/2}]\} H(t-t_{ds}) \quad (20)$$

where

$$f(\theta) = -\frac{[2(k^2-1)]^{1/2} c_s^2 \sigma_0}{\pi(k^2-2)\mu c_d^{1/2}}\left[\sin\theta - \frac{\cos\theta}{(k^2-1)^{1/2}}\right]^{-1/2}.$$

The complete time-order term in (20) is easy to derive if needed. Again here, the corresponding velocity and displacement head wavefront components can be obtained through successive time integrations of (20). The results are (20) except for a replacement of $(t-t_{ds})^{-1/2}$ there by $2(t-t_{ds})^{1/2}$ for the velocities $\dot{u}_{ds}$ and $\dot{v}_{ds}$, and by $(4/3)(t-t_{ds})^{3/2}$ for the displacements u_{ds} and v_{ds}.

It should be noted that the wavefront approximations (20) break down, having very large values near the critical value $\theta = \theta_{crx} = \sin^{-1}(1/k)$ and an infinite value at $\theta = \theta_{crx}$, through $f(\theta)$. In [11, p. 322] it is shown that for Lamb's line load problem the analogous spatial singularity in the critical region near θ_{crx} can be traded for one in time, which in turn can be removed through a convolution with a slower rising input function.

The head wavefronts for the critical region associated with the y-axis given by (11b), are easily obtained from (20) by just an interchange of $\sin\theta$ and $\cos\theta$ in $f(\theta)$. It follows, as $f(\theta)$ shows, that (20) holds on both of the quarter-plane edges.

The nature of the *two-sided equivoluminal wavefronts* has been discussed earlier. It remains to derive these events. Considering the x-axis critical region again, given by (11a), we now have the inverses, analogous to (14)

$$\begin{Bmatrix} \ddot{u}^B_{s2}(x,y,t) \\ \ddot{v}^B_{s2}(x,y,t) \end{Bmatrix} = -\frac{1}{\pi c_d} Im\left[\begin{Bmatrix} 1 \\ \zeta_{ds}/\beta'(\zeta_{ds}) \end{Bmatrix} \{\hat{U}_{sR}(\zeta_{ds}) - \hat{\mu}_\beta(\zeta_{ds})\}/g_s'(\zeta_{ds})\right] H(r/c_s - t) \quad (21)$$

and

$$\begin{Bmatrix} \ddot{u}^A_{s2}(x,y,t) \\ \ddot{v}^A_{s2}(x,y,t) \end{Bmatrix} = -\frac{1}{\pi c_d} Im\left[\begin{Bmatrix} 1 \\ \zeta_s/\beta'(\zeta_s) \end{Bmatrix} \{\hat{U}_{sR}(\zeta_s) - \hat{\mu}_\beta(\zeta_s)\}/g_s'(\zeta_s)\right] H(t - r/c_s) \quad (22)$$

where (21) and (22) represent, respectively, forms that will ultimately yield the precursor and regular parts of the acceleration two-sided equivoluminal wavefront components. We note that ζ_{ds} and ζ_s, which appear in (21) and (22), respectively, are given by (12) and (10), respectively. Superscripts B (for "before") and A(for "after") in (21) and (22) indicate the precursor and regular parts of this event. Subscript 2 indicates the two-sided nature of the event. To reduce (21) and (22) to the desired results as indicated earlier, we expand $g_s(\zeta)$ along path C away in both directions from the corner at the saddle point $-\zeta_0 = -k\sin\theta$ or $t = r/c_s$.

Consider (21) first. The time domain is $r/c_s - \epsilon \le t \le r/c_s$, which is indicated in Fig. 1 by the small portion b of path C to the right of $t = r/c_s$. Note that the earliest time here is $t = r/c_s - \epsilon$, with $0 \le \epsilon << 1$, and that time grows along b toward the saddle point $-\zeta_0$ at $t = r/c_s$. Expanding $g_s(\zeta)$ at $-\zeta_0$ along b gives

$$t = g_s(\zeta_{ds}) \simeq r/c_s + [g_s''(-\zeta_0)/2](\zeta_{ds} + \zeta_0)^2 \quad (23)$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where $g_s''(-\zeta_0) = i^2 r c_s / c_d^2 \cos^2\theta = -r c_s / c_d^2 \cos^2\theta$. From (23) we have

$$\zeta_{ds} + \zeta_0 = [2(t - r/c_s)/g_s''(-\zeta_0)]^{1/2}. \tag{24}$$

Noting, however, that on path b, $\zeta_{ds} + \zeta_0$ is a real positive vector, from (24) we must have

$$\zeta_{ds} + \zeta_0 = [2(r/c_s - t)/|g_s''(-\zeta_0)|]^{1/2}. \tag{25}$$

Then using (23) again, and (25), we have

$$g_s'(\zeta_{ds}) \simeq g''(-\zeta_0)(\zeta_{ds} + \zeta_0) = -[2|g''(-\zeta_0)|(r/c_s - t)]^{1/2}. \tag{26}$$

For (22) we proceed similarly. The time domain is $r/c_s \le t \le r/c_s + \epsilon$, $0 \le \epsilon << 1$, which is indicated in Fig. 1 by the small portion a of path C below the corner at $t = r/c_s$. With $t = r/c_s$ at the corner being the initial time in this event, time grows downward and to the left along a. Expanding $g_s(\zeta)$ at $-\zeta_0$ along a gives

$$t = g_s(\zeta_s) \simeq r/c_s + [g_s''(-\zeta_0)/2][\zeta_s + \zeta_0]^2, \tag{27}$$

and from (27),

$$\zeta_s + \zeta_0 = [2(t - r/c_s)/g_s''(-\zeta_0)]^{1/2}. \tag{28}$$

Noting, however, that on path a, $\zeta_s + \zeta_0$ is a negative imaginary vector, from (28) we must have

$$\zeta_s + \zeta_0 = -[2(t - r/c_s)/|g_s''(-\zeta_0)|]^{1/2}. \tag{29}$$

Then using (27) again, and (29), we have

$$g_s'(\zeta_s) \simeq g''(-\zeta_0)(\zeta_s + \zeta_0) = i[2|g_s''(-\zeta_0)|(t - r/c_s)]^{1/2}. \tag{30}$$

Now by approximating each of $\hat{U}_{sR}(\zeta_{ds}) - \hat{\mu}_\beta(\zeta_{ds})$ and $\hat{U}_{sR}(\zeta_s) - \hat{\mu}_\beta(\zeta_s)$ by $\hat{U}_{sR}(-\zeta_0) - \hat{\mu}_\beta(-\zeta_0)$ in (21) and (22), and making use of (26) and (30) there and that now $\alpha' = -i(k^2 \sin^2\theta - 1)^{1/2}$ and $\beta' = k \cos\theta$, (21) and (22) reduce to the acceleration two-sided equivoluminal wavefront components

$$\begin{Bmatrix} \ddot{u}_{s2}(x,y,t) \\ \ddot{v}_{s2}(x,y,t) \end{Bmatrix} = -\frac{\cos\theta}{\pi\sqrt{2c_s r}} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} Im\left\{ \begin{pmatrix} -1 \\ i \end{pmatrix} [\hat{U}_{sR}(-\zeta_0) - \hat{\mu}_\beta(-\zeta_0)] \right\} \begin{Bmatrix} (r/c_s - t)^{-1/2} H(r/c_s - t) \\ (t - r/c_s)^{-1/2} H(t - r/c_s) \end{Bmatrix} \tag{31}$$

where

$$\hat{U}_{sR}(-\zeta_0) = -\{1/R(-\zeta_0)\}\{-ik^3(1 - 2\sin^2\theta)(k^2\sin^2\theta - 1)^{1/2}\cos\theta\, \hat{\mu}^\infty_{-\alpha}(-\zeta_0) + k^3[k(1 - 2\sin^2\theta)^2 + i4\sin^2\theta\cos\theta(k^2\sin^2\theta - 1)^{1/2}]\hat{\mu}^\infty_{-\beta}(-\zeta_0) + c_d k^2 \cos\theta[i2(k^2 - 2)\sin\theta(k^2\sin^2\theta - 1)^{1/2}\hat{u} + k(1 - 2\sin^2\theta)\hat{v}],$$

$$\hat{\mu}_\beta(-\zeta_0) = (1/2k^2)\{ik\{\sin\theta[k^3\cos\theta - (k^2 - 2)]A + [\cos\theta - k\sin^2\theta]B\}[\pi c_d/(k\cos\theta - 1)]^{1/2} - c_d^2\sigma_0/\mu\},$$

and

$$\hat{\mu}^\infty_{-\alpha}(-\zeta_0) = -(\sin\theta/2k)\Big\{(\pi c_d)^{1/2}\{[ik^4\sin^2\theta(k^2\sin^2\theta - 1)^{-1/2} - (k^2 - 2)]A + k\sin\theta[i(k^2\sin^2\theta - 1)^{-1/2} - 1]B\}[1 - i(k^2\sin^2\theta - 1)^{1/2}]^{-1/2} + c_d\{[ik^4\sin^2\theta(k^2\sin^2\theta - 1)^{-1/2} - (k^2 - 2)c_d c_{R1}^{-1}]C + k\sin\theta[ic_d c_{R1}^{-1}(k^2\sin^2\theta - 1)^{-1/2} - 1]D\}[-i(k^2\sin^2\theta - 1)^{1/2} + c_d c_{R1}^{-1}]^{-1} + kc_d^2\sigma_0\sin\theta/\mu(k^2\sin^2\theta - 1)\Big\},$$

$$\hat{\mu}^\infty_{-\beta}(-\zeta_0) = (1/2k^2)\Big\{\{-k\sin\theta[k^3\cos\theta + k^2 - 2]A - k[\cos\theta + k\sin^2\theta]B\}[\pi c_d/(k\cos\theta + 1)]^{1/2} + c_d\{-k\sin\theta[k^3\cos\theta + (k^2 - 2)c_d c_{R1}^{-1}]C - k[c_d c_{R1}^{-1}\cos\theta + k\sin^2\theta]D\}[k\cos\theta + c_d c_{R1}^{-1}]^{-1} + c_d^2\sigma_0/\mu\Big\},$$

$$R(-\zeta_0) = k^4(1 - 2\sin^2\theta)^2 - i4k^3\sin^2\theta\cos\theta(k^2\sin^2\theta - 1)^{1/2}.$$

Here again the corresponding velocity and displacement two-sided wavefront components can be obtained through successive time integrations of (31). The results are (31) except for a replacement of the pair $(r/c_s - t)^{-1/2}$ and $(t - r/c_s)^{-1/2}$ there by $2(r/c_s - t)^{1/2}$ and $2(t - r/c_s)^{1/2}$ for the velocities $\dot{u}_{s2}$ and $\dot{v}_{s2}$, and by $(4/3)(r/c_s - t)^{3/2}$ and $(4/3)(t - r/c_s)^{3/2}$ for the displacements u_{s2} and v_{s2}.

The wavefront approximations (31) break down at the critical $\theta = \theta_{crx} = \sin^{-1}(1/k)$ and at the free edge surface $\theta = \pi/2$, the extremes of the present critical range. Specifically, at $\theta = \theta_{crx}$, $\ddot{u}_{S2}$, $\dot{u}_{s2}$, and u_{s2} and $\ddot{v}_{s2}$, $\dot{v}_{s2}$, and v_{s2} all become infinite due to the last term in $\hat{\mu}^\infty_{-\alpha}(-\zeta_0)$, and at $\theta = \pi/2$ they vanish due to the $\cos\theta$ multiplier in (31).

As in the case of the head wavefronts, the two-sided equivoluminal wavefronts for the critical region associated with the y-axis are easily obtained from (31) by interchanging $\sin\theta$ and $\cos\theta$ there.

Wavefront Events for the Displacements, Velocities, and Accelerations on the Quarter-Plane Edges (Surfaces)

The Rayleigh Surface Wavefronts. Of prime importance on the quarter-plane edges are the Rayleigh wavefronts. With the loaded edge ($x = 0$) Rayleigh wavefront approximations already given by (31, Part 1) for the displacements u_R, v_R, velocities $\dot{u}_R$, $\dot{v}_R$, and accelerations $\ddot{u}_R$, $\ddot{v}_R$, their behavior being, respectively, $t - y/c_{R1}$, $H(t - y/c_{R1})$, and $\delta(t - y/c_{R1})$, we need only derive the corresponding approximations for the traction-free edge ($y = 0$).

As discussed after (45, Part 1), the Rayleigh surface waves on $y = 0$ stem from the residues associated with the simple poles at $-\zeta_{R1}^{1,8}$ shown in Fig. 6, Part 1. These poles are zeros of $R(\zeta)$ in (45, Part 1). It follows, therefore, because of these zeros and since $y = 0$, that only $\hat{U}_{dR}(\zeta)$ and $\hat{U}_{sR}(\zeta)$ there are involved in the Rayleigh wavefronts, and

$$g_d(\zeta) = g_s(\zeta) = -x\zeta/c_d. \tag{32}$$

Substituting the right-hand side of (32) into (45, Part 1) then reduces the latter to

$$\begin{Bmatrix} \bar{u}(x,0,p) \\ \bar{v}(x,0,p) \end{Bmatrix} = \frac{1}{\pi c_d p^2} \int_{Br_{sL}} Im\left\{ \left[\begin{Bmatrix} 1 \\ -\alpha'/\zeta \end{Bmatrix} \hat{U}_{dR}(\zeta) + \begin{Bmatrix} 1 \\ \zeta/\beta' \end{Bmatrix} \hat{U}_{sR}(\zeta) \right] e^{px\zeta/c_d} d\zeta \right\}. \tag{33}$$

Now from (2a) and (9), we have

$$\zeta = \zeta_d(t) = \zeta_s(t) = -c_d t/x, \tag{34}$$

which are the negative ζ-values that hold along paths L_1 and L_8, shown in Fig. 3, Part 1. In Fig. 6, Part 1, the semicircular path along L_1, about the pole $-\zeta_{R1}^1$, is shown. Another such path about the pole $-\zeta_{R1}^8$ is contributed by L_8. Both are indicated by $-\zeta_{R1}^{1,8}$ in Fig. 6, Part 1.

To obtain the contributions of the Rayleigh poles at $-\zeta_{R1}^{1,8}$ we use a contour integration for the integrals in (33) over the path

$$\Gamma = C_{R+} + C_{R-} + C_{R1L}^1 + C_{R1L}^8 = Br_{SL}, \tag{35}$$

which may be seen in Fig. 6, Part 1. The contributions of C_{R+} and C_{R-} in (35) vanish according to the discussion after (45, Part 1), involving (46)–(51). It follows from (33) and (35) then, that

$$\begin{Bmatrix} \bar{u}_R(x,0,p) \\ \bar{v}_R(x,0,p) \end{Bmatrix} = \frac{1}{\pi c_d p^2} \lim_{\Delta \to 0} \left[\int_{C_{R1L}^1} Im\left\{ \begin{bmatrix} \text{integrands} \\ \text{in (33)} \end{bmatrix} d\zeta \right\} \right.$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+\int_{C^{8}_{R1L}} Im\left\{\left[\begin{array}{c}\text{integrands}\\ \text{in (33)}\end{array}\right]d\zeta\right\}\Bigg] \tag{36}$$

in which $\zeta = -\zeta^{1,8}_{R1} + \Delta e^{i\delta} = -\zeta_{R1} + \Delta e^{i\delta}$ where $\zeta_{R1} = c_d/c_{R1}$ and Δ and δ are, respectively, the modulus and argument of the complex radius vector $\zeta + \zeta_{R1}$ on paths $C^{1,8}_{R1L}$. It remains to carry out the limit indicated in (36).

We consider the C^{1}_{R1L} integral first. On L_1, $\alpha' = -i(\zeta^2 - 1)^{1/2}$ and $\beta' = -i(\zeta^2 - k^2)^{1/2}$, hence these are the values of α' and β' in the neighborhood of $\zeta = -\zeta_{R1}$. Using these values of ζ, α', and β' in the C^{1}_{R1L} integral, and obtaining the limiting value indicated in (36), we have

$$\left\{\begin{array}{c}\bar{u}_R(x,0,p)\\ \bar{v}_R(x,0,p)\end{array}\right\} = \frac{1}{\pi c_d p^2} Im\left\{ i\left[\left\{\begin{array}{c}1\\ \alpha'(-\zeta_{R1})/\zeta_{R1}\end{array}\right\}\hat{U}^{N}_{dR}(-\zeta_{R1}) + \left\{\begin{array}{c}1\\ -\zeta_{R1}/\beta'(-\zeta_{R1})\end{array}\right\}\hat{u}^{N}_{sR}(-\zeta_{R1})\right]\right\}\frac{e^{-px/c_{R1}}}{R'(-\zeta_{R1})} \tag{37}$$

where $\hat{U}^{N}_{dR}(-\zeta_{R1})$ and $\hat{U}^{N}_{sR}(-\zeta_{R1})$ are the numerators of $\hat{U}_{dR}(-\zeta_{R1})$ and $\hat{U}_{sR}(-\zeta_{R1})$, respectively, given by

$$\begin{aligned}
\hat{U}^{N}_{dR}(-\zeta_{R1}) &= 2(k^2 - 2\zeta^2_{R1})^2 \hat{\mu}^{\infty}_{-\alpha}(-\zeta_{R1}) - 4\zeta^2_{R1}(k^2 - 2\zeta^2_{R1})\hat{\mu}^{\infty}_{-\beta}(-\zeta_{R1}) - c_d\zeta_{R1}[(k^2-2)(k^2-2\zeta^2_{R1})\hat{u} + 2\zeta_{R1}\beta'(-\zeta_{R1})\hat{v}],\\
\hat{U}^{N}_{sR}(-\zeta_{R1}) &= [(k^2 - 2\zeta^2_{R1})^3/\zeta^2_{R1}]\hat{\mu}^{\infty}_{-\alpha}(-\zeta_{R1}) - 2(k^2 - 2\zeta^2_{R1})^2\hat{\mu}^{\infty}_{-\beta}(-\zeta_{R1}) + c_d\beta'(-\zeta_{R1})[2(k^2 - 2)\zeta_{R1}\alpha'(-\zeta_{R1})\hat{u} - (k^2 - 2\zeta^2_{R1})\hat{v}]
\end{aligned}$$

where

$$\begin{aligned}
\hat{\mu}^{\infty}_{-\alpha}(-\zeta_{R1}) &= (-\zeta_{R1}/2k^2)\Big\{(\pi c_{R1})^{1/2}[-\{(k^2-2)\cos(\epsilon/2) + k^2\zeta^2_{R1}(\zeta^2_{R1}-1)^{-1/2}\sin(\epsilon/2)\}\\
&\quad + i\{k^2\zeta^2_{R1}(\zeta^2_{R1}-1)^{-1/2}\cos(\epsilon/2) - (k^2-2)\sin(\epsilon/2)\}]A\\
&\quad + \zeta_{R1}(\pi c_{R1})^{1/2}[\{((\zeta^2_{R1}-1)^{-1/2}\sin(\epsilon/2) + \cos(\epsilon/2)\} - i\{((\zeta^2_{R1}-1)^{-1/2}\cos(\epsilon/2) - \sin(\epsilon/2)\}B - c_d\zeta^2_{R1}(2\zeta^2_{R1}-1)[2(k^2-1) - i\{k^2 + 2(\zeta^2_{R1}-1)\}(\zeta^2_{R1}-1)^{-1/2}]C + c_d\zeta^2_{R1}(2\zeta^2_{R1}-1)[2 - i\{(\zeta^2_{R1}-1)^{-1/2}]D - c_d\zeta_{R1}\sigma_0/\mu(\zeta^2_{R1}-1)\}\Big\},\\
\hat{\mu}^{\infty}_{-\beta}(-\zeta_{R1}) &= (1/2k^2)\Big\{(\pi c_d)^{1/2}\zeta_{R1}\{\zeta^2_{R1} - (k^2-1)\}^{-1/4}[\{(k^2-2)\cos(\omega/2) + k^2(\zeta^2_{R1}-k^2)^{1/2}\sin(\omega/2)\} - i\{k^2(\zeta^2_{R1}-k^2)^{1/2}\cos(\omega/2) - (k^2-2)\sin(\omega/2)\}]A - (\pi c_d)^{1/2}\{\zeta^2_{R1} - i(\zeta^2_{R1} - k^2)^{1/2}\}\{\zeta^2_{R1} - (k^2-1)^{-1/4}\{\cos(\omega/2) + i\sin(\omega/2)\}B + c_d\zeta_{R1}(2\zeta^2_{R1}-k^2)^{-1}[\{2(k^2-1)\zeta^2_{R1} - k^4\} - i\{2\zeta_{R1}(\zeta^2_{R1}-k^2)^{1/2}\}]C - c_d\zeta_{R1}D + c_d^2\sigma_0/\mu\Big\},\\
R'(-\zeta_{R1}) &= 8\zeta_{R1}(k^2 - 2\zeta^2_{R1}) + 4[\zeta^3_{R1}\{(\zeta^2_{R1}-k^2)^{1/2}(\zeta^2_{R1}-1)^{-1/2} + (\zeta^2_{R1}-1)^{1/2}(\zeta^2_{R1}-k^2)^{-1/2}\} + 2\zeta_{R1}\{(\zeta^2_{R1}-1)(\zeta^2_{R1}-k^2)\}^{1/2}],\\
\alpha'(-\zeta_{R1}) &= -i(\zeta^2_{R1}-1)^{1/2}, \quad \beta'(-\zeta_{R1}) = -i(\zeta^2_{R1}-k^2)^{1/2},
\end{aligned}$$

and where

$$\left[\begin{array}{c}\cos(\epsilon/2)\\ \sin(\epsilon/2)\end{array}\right] = [\zeta_{R1} \pm 1)/2\zeta_R]^{1/2}, \left[\begin{array}{c}\cos(\omega/2)\\ \sin(\omega/2)\end{array}\right] = \{[1 \pm \{\zeta^2_{R1} - (k^2-1)\}^{-1/2}]/2\}^{1/2}.$$

We now consider the C^{8}_{R1L} integral. On L_8, $\alpha' = i(\zeta^2 - 1)^{1/2}$ and $\beta' = i(\zeta^2 - k^2)^{1/2}$, hence these are the values of α' and β' in the neighborhood of $\zeta = \zeta_{R1}$. We note also that these values of α' and β' are the conjugates of those in the preceding analysis of the C^{1}_{R1L} integral which led to (37). Hence, our present case gives the conjugates of the expressions (for $\bar{u}_R$ and $\bar{v}_R$) within the inner brackets of (37) which have i as a multiplier. It follows that the sum of the two bracketed expressions lead to real results that are twice the value of the real parts of the inner bracketed expressions in (37). Therefore, multiplying the right-hand side of (37) by 2, and inverting it, gives for the acceleration Rayleigh wavefront components on the edge $y = 0$,

$$\left\{\begin{array}{c}\ddot{u}_R(x,0,t)\\ \ddot{v}_R(x,0,t)\end{array}\right\} = \frac{2}{c_d R'(-\zeta_{R1})} Im\{\text{in(37)}\}\delta(t - x/c_{R1}). \tag{38}$$

The corresponding velocity and displacement Rayleigh wavefront components on the edge $y = 0$ are obtained through successive time integrations of (38). The results are (38) except for a replacement of the $\delta(t - x/c_{R1})$ there by $H(t - x/c_{R1})$ for the velocities $\dot{u}_R$ and $\dot{v}_R$, and by $(t - x/c_{R1})H(t - x/c_{R1})$ for the displacements u_R and v_R. It should be noted that these wavefronts have the same time behavior as those on the loaded edge $x = 0$ as (31, Part 1) and the discussion following that equation point out.

The Dilatational Wavefronts. The dilational wavefronts on the unloaded edge $y = 0$ are derived from (45, Part 1). Substituting $y = 0$ into (45, Part 1) and using (32) again gives

$$\left\{\begin{array}{c}\bar{u}(x,0,p)\\ \bar{v}(x,0,p)\end{array}\right\} = \frac{1}{\pi c_d p^2}\int_{Br_{sL}} Im\left\{\left[\left\{\begin{array}{c}1\\ -\alpha'/\zeta\end{array}\right\}\left\{\hat{U}_{dR}(\zeta) + \left(\begin{array}{c}-\\ +\end{array}\right)\hat{\mu}^{\infty}_{-\alpha}(\zeta)\right\} + \left\{\begin{array}{c}1\\ \zeta/\beta'\end{array}\right\}\left\{\hat{U}_{sR}(\zeta) + \left(\begin{array}{c}-\\ +\end{array}\right)\hat{\mu}^{\infty}_{-\beta}(\zeta)\right\}\right]e^{px\zeta/c_d}d\zeta\right\}. \tag{39}$$

Now noting from Fig. 6, Part 1, that the path $-C(y = 0)$ also contains path L_2, on which $\zeta = -c_d t/x$ from (34), we see the saddle point $-\zeta_0$ has moved to the branch point $\zeta = -1$. Hence, an expansion from $\zeta = -(1+)$, along L_2 to the left in the figure, will provide the wavefront approximations we seek. So, transforming (39) with $\zeta = -c_d t/x$ and $d\zeta = -(c_d/x)dt$, and noting $\zeta = -1$ corresponds to $t = x/c_d$, (39) becomes

$$\left\{\begin{array}{c}\bar{u}_d(x,0,p)\\ \bar{v}_d(x,0,p)\end{array}\right\} = \frac{1}{\pi x p^2}\int_{x/c_d}^{\infty} Im[\text{in(39)}]_{\zeta = -(1+)} e^{-pt}dt. \tag{40}$$

Now on L_2, and hence in the neighborhood of $\zeta = -1$,

$$\begin{aligned}
\alpha'(\zeta) &= i(\zeta^2 - 1)^{1/2} \simeq [2(\zeta + 1)]^{1/2}\\
&= -i[2(c_d t/x - 1)]^{1/2},\\
\beta'(\zeta) &= i(\zeta^2 - k^2)^{1/2} \simeq (k^2 - 1)^{1/2},
\end{aligned} \tag{41}$$

in agreement with $\alpha = -i(\rho_1\rho_2)^{1/2}$, $\beta = (\rho_3\rho_4)^{1/2}$, the basic values on L_2. Substituting the values (41) for α' and β' into (40), then inverting the latter and retaining the leading wavefront terms there, the acceleration dilatational wavefront components are found to be

$$\left\{\begin{array}{c}\ddot{u}_d(x,0,t)\\ \ddot{v}_d(x,0,t)\end{array}\right\} = \frac{c_d^{3/2}\sigma_0}{\pi(k^2-2)}\left[\left\{\begin{array}{c}-[2(k^2-1)]^{1/2}/(k^2-2)\\ 2^{-1/2}\end{array}\right\}[x(t - x/c_d)]^{-1/2}H(t - x/c_d)\right] + x^{-1}0[H(t - x/c_d)]. \tag{42}$$

Again the corresponding velocity and displacement wavefront components can be obtained from (42) by successive integrations. The results are a replacement of $(t - x/c_d)^{-1/2}$ in (42) by $2(t - x/c_d)^{1/2}$ for $\dot{u}_d$, $\dot{v}_d$, and by $(4/3)(t - x/c_d)^{3/2}$ for u_d, v_d.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Equivoluminal Wavefronts. It was pointed out in the discussion after the head wavefront approximations for the interior (20), that these hold also on the edges of the quarter plane. Hence, here we need only derive the *two-sided equivoluminal wavefronts* for the edges of the quarter plane. We consider the $y=0$ edge first, where (39) holds, and note that the saddle pont $-\zeta_0=-k\sin\theta$ has moved to the branch point $\zeta=-k$. Hence, expansions from $\zeta=-(k+)$ to the left along path L_1, and from $\zeta=-(k-)$ to the right along path L_2 will provide the wavefront approximations we seek. We again have $\zeta=-c_d t/x$ and $d\zeta=-(c_d/x)dt$, and $\zeta=-k$ corresponds to $t=x/c_s$. Along L_2 in the neighborhood of $\zeta=-k$ then, (39) becomes

$$\begin{Bmatrix} \ddot{u}^B_{s2}(x,0,p) \\ \ddot{v}^B_{s2}(x,0,p) \end{Bmatrix} = \frac{1}{\pi x p^2}\int_{x/c_s-\epsilon}^{x/c_S} Im[\text{in}(39)]_{\zeta=-(k-)}e^{-pt}dt, \quad (43)$$

where $0\le\epsilon<<1$, and along L_1 in the neighborhood of $\zeta=-k$,

$$\begin{Bmatrix} \ddot{u}^A_{s2}(x,0,p) \\ \ddot{v}^A_{s2}(x,0,p) \end{Bmatrix} = \frac{1}{\pi x p^2}\int_{x/c_s}^{\infty} Im[\text{in}(39)]_{\zeta=-(k+)}e^{-pt}dt. \quad (44)$$

Again as in (21) and (22) for the interior, the superscripts B (for before) and A (for after) in (43) and (44) indicate the precursor and regular parts of this two-sided event.

Treating the precursor (43) first, on L_2, near $\zeta=-(k-)$,

$$\alpha'=i(\zeta^2-1)^{1/2}=-i(k^2-1)^{1/2},$$
$$\beta'=i(\zeta^2-k^2)^{1/2}=i[-2k(\zeta+k)]^{1/2}=k\{2[1-(c_s t/x)]\}^{1/2}, \quad (45)$$

in agreement with the basic values $\alpha=-i(\rho_1\rho_2)^{1/2}$ and $\beta=(\rho_3\rho_4)^{1/2}$ on L_2. Then substituting the values (45) for α' and β' into (43), inverting the latter and retaining the leading wavefront terms there, one finds the precursor parts of the acceleration two-sided equivoluminal wavefront components to be

$$\begin{Bmatrix} \ddot{u}^B_{s2}(x,0,t) \\ \ddot{v}^B_{s2}(x,0,t) \end{Bmatrix} = \frac{(k^2-1)^{1/2}}{\pi k x}\left[\begin{Bmatrix} 2(2c_s/x)^{1/2}/k^2 \\ -1 \end{Bmatrix}\right.$$
$$\left.\begin{Bmatrix} E^B_u(x/c_s-t)^{1/2} \\ E^B_{v1}+E^B_{v2}(x/c_s-t)^{1/2} \end{Bmatrix}\right]H(x/c_s-t) \quad (46)$$

where

$$E^B_u=2(2k^2+1)Re\hat{\mu}^{\infty}_{-\alpha}(-k)+c_d k(k^2-2)\hat{u},$$
$$E^B_{v1}=-\{2Re\hat{\mu}^{\infty}_{-\alpha}(-k)+4\hat{\mu}^{\infty}_{-\beta l}[-(k-)]+c_d(k^2-2)\hat{u}/k\},$$
$$E^B_{v2}=-(2\sqrt{2c_s}c_d/k)\hat{v},$$

and where

$$Re\hat{\mu}^{\infty}_{-\alpha}(-k)=-(2k)^{-1}\left\{(\pi c_s)^{1/2}\{[-(k^2-2)\cos q + k^4(k^2-1)^{-1/2}\sin q]A+k[-\cos q+(k^2-1)^{-1/2}\sin q]B\}\right.$$
$$\left.-c_d(c_d^2/c_{R1}^2+k^2-1)^{-1}\{[(k^2-2)c_d^2/c_{R1}^2+k^4]C+(2kc_d/c_{R1})D\}+\frac{c_d^2 k\sigma_0}{\mu(k^2-1)}\right\},$$
$$\hat{\mu}^{\infty}_{-\beta l}[-(k-)]\simeq-(2k)^{-1}[(c_d\pi)^{1/2}\{(k^2-2)A+kB\}+c_d(k^2-2)C+kc_{R1}D-c_d^2\sigma_0/k\mu],$$

and $q=\tan^{-1}[-(k^2-1)^{1/2}/2]$. The subscript l on $\hat{\mu}^{\infty}_{-\beta l}[-(k-)]$ indicates that $\hat{\mu}^{\infty}_{-\beta}$ here represents the coefficient of its leading wavefront term.

It remains to treat (44) similarly which will yield the regular wavefronts for the two-sided equivoluminal events. Here on L_1, near $\zeta=-(k+)$, α' does not change from its value on L_2 given by the first of (45). However, β' changes its nature, i.e.,

$$\beta'=i(\zeta^2-k^2)^2=[2k(\zeta+k)]^{1/2}=-ik[2(c_s t/x-1)]^{1/2}. \quad (47)$$

These values of α' and β' are in agreement with the basic values $\alpha=-i(\rho_1\rho_2)^{1/2}$ and $\beta=-i(\rho_3\rho_4)^{1/2}$ on L_1. Therefore, substituting the right-hand sides of the first of (45) and (47) for α' and β' into (44), inverting the latter, and retaining the leading wavefronts terms there, one finds the regular parts of the acceleration two-sided equivoluminal wavefront components to be

$$\begin{Bmatrix} \ddot{u}^A_{s2}(x,0,t) \\ \ddot{v}^A_{s2}(x,0,t) \end{Bmatrix} = \frac{2(k^2-1)}{\pi k x}\left[\begin{Bmatrix} 2k^{-2}(2c_s x)^{-1/2} \\ 1 \end{Bmatrix}\right.$$
$$\left.\begin{Bmatrix} E^A_u(t-x/c_s)^{1/2} \\ E^A_v \end{Bmatrix}\right]H(t-x/c_s) \quad (48)$$

where

$$E^A_u=(2k^2+1)Im\hat{\mu}^{\infty}_{-\alpha}(-k)+Im\hat{\mu}_{-\beta l}[-(k+)]+3c_d k^3\hat{v}$$
$$E^A_v=Re\hat{\mu}^{\infty}_{-\alpha}(-k)-2\text{Re}\hat{\mu}^{\infty}_{-\beta l}[-(k+)]-c_d(k^2-2)[(k^2-1)^{1/2}-2k^{-1}]\hat{u}$$

and where

$$Im\hat{\mu}^{\infty}_{-\alpha}(-k)=-(2k)^{-1}\left\{(\pi c_s)^{1/2}\{[k^4(k^2-1)^{-1/2}\cos q + (k^2-2)\sin q]A+k[(k^2-1)^{-1/2}\cos q+\sin q]B\}\right.$$
$$\left.+c_d(c_d^2/c_{R1}^2+k^2-1)^{-1}\{(c_d/c_{R1})(k^4(k^2-1)-(k^2-2)(k^2-1)^{1/2}]C+k[(c_d^2/c_{R1}^2)(k^2-1)^{-1/2}-(k^2-1)^{1/2}]D\}\right\}$$
$$Im\hat{\mu}^{\infty}_{-\beta l}[-(k+)]=-(2k^2)^{-1}\left\{k(\pi c_s)^{1/2}\{k[k+(k^2-2)/2]A-(1-k^2/2)B\}-k^2 c_{R1}\{2(k^2-1)C+(c_d/c_{R1}-kc_{R1}/c_d)D\}\right\}$$

and where $Re\ \hat{\mu}^{\infty}_{-\alpha}(-k)$ and q are given in (46), while $Re\ \hat{\mu}^{\infty}_{-\beta l}[-(k+)]$ is equal to $\hat{\mu}^{\infty}_{-\beta l}[-(k-)]$ in (46).

The complete acceleration two-sided equivoluminal wavefront components $\ddot{u}_{s2}(x,0,t)$ and $\ddot{v}_{s2}(x,0,t)$ are formed by each of the pairs $\ddot{u}_{s2}$ and $\ddot{v}_{s2}$ obtained from (46) and (48). Note the time behavior of the pairs u^B_{s2} and u^A_{s2} are the same, i.e., $(x/c_s-t)^{1/2}$ and $(t-x/c_s)^{1/2}$. Similarly, v^B_{s2} and v^A_{s2} are $H(x/c_s-t)$ and $H(t-x/c_s)$, respectively. Important is the fact that these $y=0$ edge (surface) wavefronts are much less singular than the corresponding interior wavefronts given by (31). This type of behavior is well known in Lamb's problem [cf. 11, Figs. 6.14, 6.15, and 6.16].

As in our earlier wavefront events, the velocities $\dot{u}_{s2}(x,0,t)$ and $\dot{v}_{s2}(x,0,t)$ and the displacements $u_{s2}(x,0,t)$ and $v_{s2}(x,0,t)$ are obtained from (46) and (48) by successive integrations.

Consider now the *two-sided equivoluminal wavefronts on the loaded edge* $x=0$. From (45, Part 1) these events are governed by

$$\begin{Bmatrix} \bar{u}_{s2}(0,y,p) \\ \bar{v}_{s2}(0,y,p) \end{Bmatrix} = \frac{1}{\pi c_d p^2}\int_{Br_{sL}} Im\left[\begin{Bmatrix} 1 \\ \zeta/\beta'(\zeta) \end{Bmatrix}\right.$$
$$\left.\{\hat{U}_{sR}(\zeta)-\hat{\mu}_{\beta}(\zeta)\}e^{-pg_s(\zeta)}d\zeta\right] \quad (49)$$

where $\hat{U}_{sR}$, $\hat{\mu}_{\beta}$, $R(\zeta)$, β', and α' are defined in (45, Part 1). Now in the critical region $0<\theta<\theta_{\text{cry}}$, we have for the y-axis ($x=0$) edge from (9) and (10),

$$g_s(\zeta)=\beta'(\zeta)y/c_d, \quad \zeta_s(t)=-ik[(c_s t/y)^2-1]^{1/2}, \quad (50)$$

respectively. From the second of (50) we see that our ζ-contour includes the lower half of the $Im\ \zeta$-axis, the origin and the positive $Re\ \zeta$ part of path L_3, cf. Fig. 1. This equation also shows that now the origin corresponds to the time

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$t=y/c_s$. The next step is a contour integration involving the ζ-contour shown in Fig. 1. We find, using the first of (50), that (49) can be traded for

$$\left\{\begin{matrix}\bar{u}_{s2}(0,y,p)\\ \bar{v}_{s2}(0,y,p)\end{matrix}\right\}=\frac{1}{\pi c_d^2 p^2}\left\{\left[\int_{-i\infty}^{0}+\int_{L_3}\right]Im\left[\left\{\begin{matrix}1\\ \zeta/\beta'(\zeta)\end{matrix}\right\}\right.\right.$$

$$\left.\left.\{\hat{U}_{sR}(\zeta)-\hat{\mu}_\beta(\zeta)\}e^{-p\beta' y/c_d}d\zeta\right]\right\} \tag{51}$$

Now from the first of (50),

$$g_s(\zeta)=t=\beta' y/c_d;\quad \beta'=c_d t/y. \tag{52}$$

β' in (52) is positive, in agreement with its basic values $\beta=(\rho_3\rho_4)^{1/2}$ on the lower half $Im\ \zeta$-axis and L_3. α' also has the same character there, i.e., $\alpha=(\rho_1\rho_2)^{1/2}$. From the second of (50) we see the wavefronts here stem from small ζ. Indeed, the paths b' and a' near the origin of ζ in Fig. 1 will be shown to yield from (51) the precursors and regular parts, respectively, of the two-sided equivoluminal wavefronts. We treat first the precursors from the L_3 integral in (51), noting on path b', with small ζ, that

$$\left\{\begin{matrix}\bar{u}^B_{s2}(0,y,p)\\ \bar{v}^B_{s2}(0,y,p)\end{matrix}\right\}=-\frac{1}{\pi c_d p^2}\int_\epsilon^0$$

$$Im\left[\left\{\begin{matrix}1\\ \zeta_s(t)/\beta'[\zeta_s(t)]\end{matrix}\right\}[\hat{U}_{sR}[\zeta_s(t)]\right.$$

$$\left.-\hat{\mu}_\beta[\zeta_s(t)]\}e^{-p\beta' y/c_d}d\zeta_s(t)\right] \tag{53}$$

where $0\le\epsilon<<1$.

Compatible with the small value of $\zeta_s(t)$ indicated in (53) are the relations

$$\begin{aligned}&\alpha'=i(\zeta^2-1)\simeq i(-i)=1,\quad \beta'=c_d t/y\simeq k,\\ &\zeta_s(t)=k[1-(c_s t/y)^2]^{1/2}\simeq k[2(1-c_s t/y)]^{1/2},\\ &d\zeta_s(t)=-kc_s^2 y^{-2}[1-(c_s t/y)^2]^{-1/2}dt,\\ &(1-\beta')^{1/2}=[i^2(k-1)]^{1/2}=i(k-1)^{1/2}.\end{aligned} \tag{54}$$

From (52) and (54) we then have from (53),

$$\left\{\begin{matrix}\bar{u}^B_{s2}(0,y,p)\\ \bar{v}^B_{s2}(0,y,p)\end{matrix}\right\}=\frac{c_s^{1/2}}{2[2\pi c_d(k-1)y]^{1/2}p^2}\int_{y/c_s-\delta}^{y/c_s}$$

$$Im\left[i\left\{\begin{matrix}S^B_{u1}(y/c_s-t)^{-1/2}+y^{-1/2}+y^{-1/2}S^B_{u2}\\ y^{-1/2}S^B_{v1}+y^{-1}S^B_{v2}(y/c_s-t)^{1/2}\end{matrix}\right\}\right]e^{-pt}dt \tag{55}$$

where

$$S^B_{u1}=-B,\quad S^B_{u2}=(2c_s)^{1/2}(k^3-k^2+2)A,$$

$$S^B_{v1}=-(2c_s)^{1/2}B,\quad S^B_{v2}=c_s(k^3-k^2+2)A,$$
$$\delta=\epsilon^2 y/2c_s k^2.$$

From (55) then, one finds the leading wavefront terms for the precursor parts of the two-sided equivoluminal acceleration components to be

$$\left\{\begin{matrix}\ddot{u}^B_{s2}(0,y,t)\\ \ddot{v}^B_{s2}(0,y,t)\end{matrix}\right\}=\frac{c_s^{1/2}}{2[2\pi c_d(k-1)y]^{1/2}}$$

$$\left\{\begin{matrix}S^B_{u1}(y/c_s-t)^{-1/2}+y^{-1/2}S^B_{u2}\\ y^{-1/2}S^B_{v1}+y^{-1}S^B_{v2}(y/c_s-t)^{1/2}\end{matrix}\right\}H(y/c_s-t) \tag{56}$$

where $y/c_s-\delta\le t\le y/c_s$.

The regular parts of the two-sided equivoluminal wavefronts are obtained from the first integral (along the negative $Im\ \zeta$-axis) in (51), noting on path a', with small ζ, that

$$\left\{\begin{matrix}\bar{u}^A_{s2}(0,y,p)\\ \bar{v}^A_{s2}(0,y,p)\end{matrix}\right\}=-\frac{1}{\pi c_d p^2}\int_0^{-i\infty}$$

$$Im\left[\left\{\begin{matrix}1\\ \zeta_s(t)/\beta'[\zeta_s(t)]\end{matrix}\right\}\{\hat{U}_{sR}[\zeta_s(t)]\right.$$

$$\left.-\hat{\mu}_\beta[\zeta_s(t)]\}e^{-p\beta' y/c_d}d\zeta_s(t)\right]. \tag{57}$$

Now compatible with small values of $\zeta_s(t)$ in (57) are the relations for α', β', and $(1-\beta')^{1/2}$ in (54), and

$$\begin{aligned}\zeta_s(t)&=-ik[(c_s t/y)^2-1]^{1/2}\simeq -ik[2(c_s t/y-1)]^{1/2},\\ d\zeta_s(t)&=-ikc_s y^{-1}[(c_s t/y)^2-1]^{-1/2}.\end{aligned} \tag{58}$$

Using (52), the values for α', β', and $(1-\beta')^{1/2}$ in (54) and the relations (58), we then have from (57),

$$\left\{\begin{matrix}\bar{u}^A_{s2}(0,y,p)\\ \bar{v}^A_{s2}(0,y,p)\end{matrix}\right\}=-\frac{c_s^{1/2}}{\pi c_d(2y)^{1/2}p^2}\int_{y/c_s}^{\infty}$$

$$Im\left[i\left\{\begin{matrix}S^A_{u1}(t-y/c_s)^{-1/2}+y^{-1/2}S^A_{u2}\\ y^{-1/2}S^A_{v1}+y^{-1}S^A_{v2}(t-y/c_s)^{1/2}\end{matrix}\right\}\right]e^{-pt}dt \tag{59}$$

where

$$\begin{aligned}S^A_{u1}&=-2^{-1/2}[(\pi c_d/(k+1))^{1/2}B-2c_d^3 c_{R1}^{-2}(k^2\\ &\quad -c_d^2 c_{R1}^{-2})^{-1}D-c_d\hat{v}],\\ S^A_{u2}&=-(c_s/2)^{1/2}[k^3-(k^2-2)][\pi c_d/(k-1)]^{1/2}A,\\ S^A_{v1}&=-[\pi c_d c_s/2(k-1)]^{1/2}B,\\ S^A_{v2}&=2c_s\left\{\{(\pi c_d)^{1/2}[2^{1/2}(k^2-2)/k-(k^3+k^2\right.\\ &\quad -2)/2(1+k)^{1/2}]\}A+\{2(k^2-2)c_d^2 c_{R1}^{-1}/k(1\\ &\quad +c_d c_{R1}^{-1})+c_d[k^3+(k^2-2)c_d c_{R1}^{-1}]/[k+c_d c_{R1}^{-1}]\\ &\quad -(c_d/2)[k^3-(k^2-2)c_d c_{R1}^{-1}]/[k-c_d c_{R1}^{-1}]\}C\\ &\quad \left.-2c_d(k^2-2)\hat{u}/k\right\}.\end{aligned}$$

From (59) then, one finds the leading wavefront terms for the regular parts of the two-sided equivoluminal acceleration components to be

$$\left\{\begin{matrix}\ddot{u}^A_{s2}(0,y,t)\\ \ddot{v}^A_{s2}(0,y,t)\end{matrix}\right\}=-\frac{c_s^{1/2}}{\pi c_d(2y)^{1/2}}$$

$$\left\{\begin{matrix}S^A_{u1}(t-y/c_s)^{-1/2}+y^{-1/2}S^A_{u2}\\ y^{-1/2}S^A_{v1}+y^{-1}S^A_{v2}(t-y/c_s)\end{matrix}\right\}H(t-y/c_s) \tag{60}$$

where $y/c_s\le t\le y/c_s+\epsilon$, with $0\le\epsilon<<1$.

The complete acceleration two-sided equivoluminal wavefront components $\ddot{u}_{s2}(0,y,t)$ and $\ddot{v}_{s2}(0,y,t)$ are formed by each of the pairs $\ddot{u}_{s2}$ and $\ddot{v}_{s2}$, obtained from (56) and (60). Note again here the time behavior of the pairs $\ddot{u}^B_{s2}$ and $\ddot{u}^A_{s2}$ are the same, i.e., $(y/c_s-t)^{-1/2}$ and $(t-y/c_s)^{-1/2}$. Similarly, $\ddot{v}^B_{s2}$ and $\ddot{v}^A_{s2}$ are $H(y/c_s-t)$ and $H(t-y/c_s)$, respectively. Important here is the fact that the $\ddot{u}_{s2}$ wavefront singularities are the same as those of corresponding interior results given in (31). The same is not true of the $\ddot{v}_{s2}$ wavefronts which we have already noted behave as H functions, whereas the corresponding $\ddot{v}_{s2}$ wavefronts in (31) have the infinite singularities $(y/c_s-t)^{-1/2}$ and $(t-y/c_s)^{-1/2}$. Further, it is interesting to note that the $\ddot{u}_{s2}(0,y,t)$ singularities match the singularities of the dilatational wavefronts given by (28). Also observe that the $\ddot{u}_{s2}(0,y,t)$ wavefronts are much more singular than those of $\ddot{u}_{s2}(x,0,t)$. Perhaps the *most important comparison of the two-sided equivoluminal wavefront singularities* stems from the symmetries of these events relative to the x and y-axes. That is, along the x-axis, $\ddot{u}_{s2}(x,0,t)\sim(x/c_s-t)^{1/2}$ and $(t-x/c_s)^{1/2}$, whereas along the y-axis its mate, $\ddot{v}_{s2}(0,y,t)\sim H(y/c_s-t)$ and $H(t-y/c_s)$. Similarly, the normal acceleration along the x-axis, $\ddot{v}_{s2}(x,0,t)\sim H(x/c_s-t)$ and $H(t-x/c_s)$, whereas the normal acceleration along the y-axis, $\ddot{u}_{s2}(0,y,t)\sim(y/c_s-t)^{-1/2}$ and $(t-y/c_s)^{-1/2}$. Clearly, these wavefronts show the stronger singularities along the loaded edge $x=0$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Finally, it is important to note that our use of the positive $Re\ \zeta$ part of path L_3 in the preceding analysis for the $x=0$ edge does not violate unboundedness via the reasoning leading to the vanishing of L_3 in Fig. 5, Part 1. Clearly, this is so, since with $x=0$, $\exp(sx) = \exp[k_d Re(\zeta)x] = 1$ along L_3, hence we have boundedness there.

The Dilatational Plane Wavefront. To obtain this wavefront we return to (45, Part 1) and draw on the terms there exhibiting the singularity $(\zeta+1)^{-1}$ at $\zeta=-1$, shown in Fig. 6, Part 1. The terms that contribute are contained in

$$\bar{u}_{dp}(x,y,p) = \frac{2}{\pi c_d p^2}\left[\int_{Br_{sL}} Im[-\hat{\mu}^{\infty}_{-\alpha}(\zeta)e^{k_d\zeta x}d\zeta] - \int_{Br_{sU}} Im[-\hat{\mu}^{\infty}_{-\bar{\alpha}}(\zeta)e^{k_d\bar{\zeta}x}d\bar{\zeta}]\right] \tag{61a}$$

where

$$-\hat{\mu}^{\infty}_{-\alpha}(\zeta) = c_d^2\sigma_0\zeta^2/2k^2\mu(\zeta^2-1),$$
$$-\hat{\mu}^{\infty}_{-\bar{\alpha}}(\bar{\zeta}) = c_d^2\sigma_0\bar{\zeta}^2/2k^2\mu(\bar{\zeta}^2-1). \tag{61b}$$

The contributions follow from contour integrations for each of the integrals in (61*a*), using (61*b*). It can be seen from (61*b*) that the contour integrals replacing (61*a*) are the lower half circular path about $\zeta=-1$, C_{-1L} for the first integral in (61*a*) (cf. Fig. 6, Part 1), and the upper half circular path C_{-1U} (not shown in Fig. 6, Part 1), for the second integral there. The second integral in (61*a*) over Br_{sU} is an alternative representation to that of Br_{sL}, useful here because of the second of (61*b*). Indeed, because of the simple pole nature of both terms in (61*b*), $(\zeta+1)^{-1}$ and $(\bar{\zeta}+1)^{-1}$, we are able to integrate fully around the circular path about $\zeta=-1$ for each of these terms. This evidences itself in (61*a*) by the multiplicative factor 2 out front in (61*a*). Mathematically we have then,

$$\bar{u}_{dp}(x,y,p) = \frac{c_d\sigma_0}{\pi k^2 p^2}\left[\lim_{\zeta\to -1}\left\{\int_{C_{-1L}} Im\left(\frac{\zeta^2}{\zeta^2-1}e^{k_d\zeta x}d\zeta\right) - \int_{C_{-1U}} Im\left(\frac{\bar{\zeta}^2}{\bar{\zeta}^2-1}e^{k_d\bar{\zeta}x}d\bar{\zeta}\right)\right\}\right] = -\frac{c_d\sigma_0 e^{-k_d x}}{2\pi k^2\mu p^2}$$
$$\left[\int_{C_{-1L}} Im\left(\frac{d\zeta}{\zeta+1}\right) - \int_{C_{-1U}} Im\left(\frac{d\bar{\zeta}}{\bar{\zeta}+1}\right)\right]. \tag{62}$$

Then, substituting $\zeta+1=\Delta e^{i\eta}$ and $\bar{\zeta}+1=\Delta e^{-i\eta}$ into the first and second integrals of (62), respectively, and carrying out the integrations (in which moduli Δ on $C_{-1L,U}\to 0$) and inversion, gives the complete plane displacement wave

$$u_{dp}(x,y,t) = -(c_d\sigma_0/k^2\mu)(t-x/c_d)H(t-x/c_d), \tag{63}$$

which also serves to represent the wavefront. It follows by successive time differentiations of (63) that $\dot{u}_{dp}$ and $\ddot{u}_{dp}$ are, respectively, represented by $H(t-x/c_d)$ and $\delta(t-x/c_d)$. From the first of (2, Part 1), and the fact that v does not contribute to the present event, (63) yields the stress

$$\sigma_{xdp}(x,y,t) = (\lambda+2\mu)u_{dpx}(x,y,t) = \sigma_0 H(t-x/c_d). \tag{64}$$

It follow from (64) that on the loaded edge $x=0$, $\sigma_x(0,y,t)=\sigma_0 H(t)$, in agreement with the first of (10) in Part 1.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

D. B. Goetschel
Assistant Professor,
Department of Mechanical Engineering
Aeronautical Engineering and Mechanics,
Rensselaer Polytechnic Institute,
Troy, N.Y. 12181

S. B. Dong
Professor.
Mem. ASME

R. Muki
Professor.
Mem. ASME

Mechanics and Structures Department,
University of California,
Los Angeles, Calif. 90024

A Global Local Finite Element Analysis of Axisymmetric Scattering of Elastic Waves

A global local finite element method is developed for axisymmetric scattering of a steady, compressive, incident elastic wave in a homogeneous, isotropic host medium by an axisymmetric inclusion. The inclusion may be arbitrary with respect to its geometry and can have inhomogeneous, anisotropic elastic material properties. Examples on spheroidal and finite circular cylindrical inclusions are given. Comparison of current data with available results show good agreement.

Introduction

The mathematical treatment of steady-state elastic wave diffraction and scattering began with the works of Clebsch (1867) and Lamb (1900). Save Sezawa's notable contribution (1927), virtually no activity occurred in the first half of this century. Pao and Mow [1] provided an excellent history of this subject, which also serves as a comprehensive state-of-the-art review up to 1973. The renewed interest in such problems stems from a host of current applications which, to mention a few, include nondestructive evaluation, dynamic soil-structure interaction, and solid-state physics.

Prior to 1970, only two analytical solution techniques for steady-state three-dimensional problems are worthy of mention, and they met with minimal success as measured by the number of problems solved. One is separation of variables (also called the wave function expansion method), which applied only to the case of an isotropic spherical inclusion embedded in an isotropic medium. The other technique is the integral equation method, which is based on the mathematical interpretation of Huygen's principle. Although this method is conceptually valid for scatterers of arbitrary geometry, again analytical solutions do not appear possible except for the sphere.

In recent literature, a numerical solution method called the transition matrix appeared that essentially implements the integral equation technique. The method was proposed in various forms by Waterman [2, 3], Varatharajulu and Pao [4], and Pao [5, 6]. It uses spherical wave function in the evaluation of the interior and exterior Helmholtz integral formulas for an arbitrary shaped inclusion. By imposing displacement and traction continuity in terms of the incident, scattered and refracted wave functions, a matrix equation results for the amplitudes of the scattered wave functions in terms of those for the incident wave. Some applications on spheroidal inclusions have appeared [7]. Visscher [8, 9] raised some questions concerning: (1) the accuracy and convergence of the transition matrix (by pointing out that these items were not fully addressed); and (2) the numerical conditioning in the inverse of the so-called "Q" matrix, in an intermediate step in this method. He proposed another method called MOOT for method of optimal truncation. In MOOT, expressions are developed for displacements and tractions for both interior and exterior regions from a complete set of wave functions. Then, interface displacement continuity and traction reciprocity are enforced on the basis of a least square error minimization. It should be pointed out that MOOT relies on an ad hoc coefficient in the least square error combination of the tractions and displacements.

Although the transition matrix and MOOT are quite general, they still cannot be easily applied to inhomogeneous, anisotropic scatterers. Herein, a more versatile technique is proposed, which is an adaptation of the global-local finite element method (GLFEM). Earlier finite element analyses of problems involving a far field resulted in cumbersome and approximate models of the outer field. Moreover, the "quiet" boundary at the mesh interface to the outer field in dynamic problems invariably raised serious issues that cast doubts in the efficacy of most methods to date. GLFEM obviates many of these difficulties and has been successfully applied to a number of elastostatic half-space problems [10–12]. For a state-of-the-art survey of global-local finite element methods, see reference [13]. At present, the specific problem considered is that of axisymmetric diffraction and scattering by an inclusion with arbitrary, axisymmetric geometry and elastic properties upon illumination by a steady monochromatic

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August, 1981; final revision, January, 1982. Paper No. 82-WA/APM-3.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

compressional wave in an isotropic homogeneous host medium. The inclusion and some portion of the surrounding medium with its outer boundary in a configuration of a spherical surface undergo finite element discretization. Adopting a spherical mesh boundary enables the scattered field beyond the finite element model to be conveniently represented by spherical wave functions. Illumination by shear waves can also be treated by this discretization, however, this topic is deferred. In its full ramifications, the method is quite general and can treat arbitrary three-dimensional anisotropic, inhomogeneous scatterers.

Preliminaries

Consider a homogeneous, isotropic medium with Lamé constants λ and μ and density ρ. Embedded in it is an axisymmetric body which may be inhomogeneous and may possess cylindrical anisotropic properties. Let a steady monochromatic compressional wave shine on this inclusion. This wave's time dependence is $e^{-i\omega t}$ where ω is the frequency and its propagation direction parallels the axis of symmetry of the inclusion. The time factor $e^{-i\omega t}$ occurs throughout and may be omitted without confusion.

The proposed method utilizes a finite element model in the form of a spherical region of radius R that contains the inclusion and some portion of the surrounding medium (see Fig. 1). In this method, a forced vibration of this discrete model is perceived, where the incident and scattered waves of the outer field pose as the forcing functions acting on the finite element boundary. The governing equations of the finite element model for forced vibration have the form

$$[\mathbf{K}-\omega^2\mathbf{M}]\ \mathbf{U}=\mathbf{F} \tag{1}$$

where the stiffness $\mathbf{K}$ and consistent mass $\mathbf{M}$ and consistent load vector $\mathbf{F}$ are formulated using conventional axisymmetric finite element methodology in cylindrical coordinates. The displacement vector $\mathbf{U}$ is the array of radial and axial displacement components at the nodes. Note that $\mathbf{K}$ is singular, because no boundary conditions have been imposed to preclude rigid body motions, but $[\mathbf{K}-\omega^2\mathbf{M}]$ is not, provided ω is nonzero and does not coincide to a natural frequency of the finite element model.[1] For convenience, define $\mathbf{D}(\omega)$ as

$$\mathbf{D}(\omega)=[\mathbf{K}-\omega^2\mathbf{M}] \tag{2}$$

so that equation (1) takes the form:

$$\mathbf{D}(\omega)\mathbf{U}=\mathbf{F} \tag{3}$$

Because the forcing functions act only on the boundary nodes, equation (3) is partitioned to read:

$$\begin{bmatrix}\mathbf{D}_{ii} & \mathbf{D}_{ib}\\ \mathbf{D}_{bi} & \mathbf{D}_{bb}\end{bmatrix}\begin{bmatrix}\mathbf{U}_i\\ \mathbf{U}_b\end{bmatrix}=\begin{bmatrix}0\\ \mathbf{F}_b\end{bmatrix} \tag{4}$$

where subscripts i and b denote the interior and boundary degrees of freedom, respectively. This finite element forced vibration problem will be called the *interior problem*.

The axisymmetric wave functions exterior to the finite element model will be expressed in spherical coordinates (r,θ). In GLFEM, these functions are called the global functions and they will also be referred to as the *exterior functions*. For the scattered field, spherical harmonics that satisfy the axisymmetric wave equations for a homogeneous, isotropic medium and possess the outgoing properties will be used as the global functions. Their explicit forms may be found in Pao and Mow [1]. In the application of the method, the global functions evaluated at the mesh boundary $(r=R)$ are needed. The displacements and surface stresses at the mesh boundary have the forms:

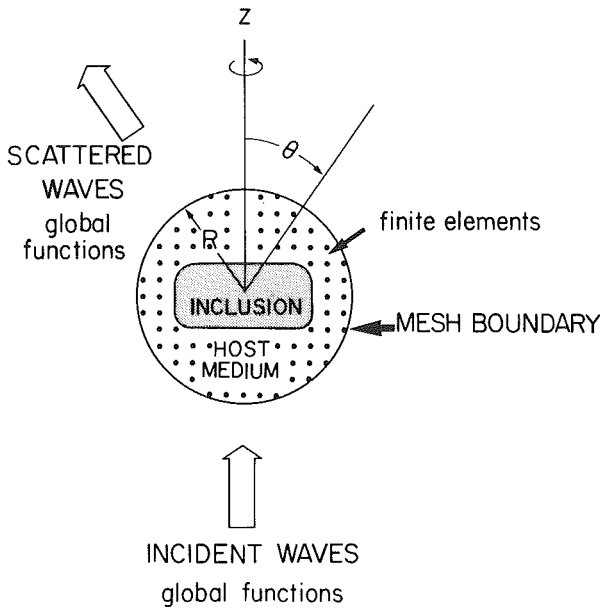


Fig. 1 Problem illustration

$$u_r(\theta)=\sum_{n=0}^{N}\left[A_nD_{1n}-B_nS_{1n}\right]p_n(\cos\theta) \tag{5}$$

$$u_\theta(\theta)=\sum_{n=0}^{N}\left[A_nD_{2n}-B_nS_{2n}\right]\frac{dP_n(\cos\theta)}{d\theta} \tag{6}$$

$$\sigma_{rr}(\theta)=\sum_{n=0}^{N}\left[A_nD_{3n}-B_nS_{3n}\right]P_n(\cos\theta) \tag{7}$$

$$\sigma_{r\theta}(\theta)=\sum_{n=0}^{N}\left[A_nD_{4n}-B_nS_{4n}\right]\frac{dP_n(\cos\theta)}{d\theta} \tag{8}$$

where A_n, B_n are complex constants denoting the strength of the particular global function, $P_n(\cos\theta)$ are Legendre polynomials, and the coefficients D_{jn} and S_{jn} are defined by:[2]

$$\begin{aligned}
D_{1n}&=\frac{n}{R}h_n(\alpha R)-\alpha h_{n+1}(\alpha R)\\
D_{2n}&=\frac{1}{R}h_n(\alpha R)\\
D_{3n}&=\frac{2\mu}{R^2}\left[(n^2-n-\tfrac{1}{2}\alpha^2R^2)h_n(\alpha R)+2\alpha R h_{n+1}(\alpha R)\right] \qquad (9)\\
D_{4n}&=\frac{2\mu}{R^2}\left[(n-1)h_n(\alpha R)-\alpha R h_{n+1}(\alpha R)\right]\\
S_{1n}&=\frac{n(n+1)}{R}h_n(\beta R)\\
S_{2n}&=\frac{(n+1)}{R}h_n(\beta R)+\beta h_{n+1}(\beta R) \qquad (10)\\
S_{3n}&=\frac{2\mu}{R^2}n(n+1)\left[(n-1)h_n(\beta R)-\beta R h_{n+1}(\beta R)\right]\\
S_{4n}&=\frac{2\mu}{R^2}\left[(n^2-n-\tfrac{1}{2}\beta^2R^2)h_n(\beta R)-\beta R h_{n+1}(\beta R)\right]
\end{aligned}$$

where h_n are spherical Bessel functions of the first kind, and

$$\alpha^2=\omega^2\rho/(\lambda+2\mu);\quad \beta^2=\omega^2\rho/\mu \tag{11}$$

[1] Coincidence of ω with a natural frequency of the finite element model can easily be avoided by changing R slightly.

[2] The coefficient B_0 is immaterial, as $S_{10}=S_{30}=0$ and $dP_0(\cos\theta)/d\theta=0$. It is included in the summation for notational convenience.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The incident field may be any axisymmetric compressional wave form whose propagation direction coincides with the inclusion's axis of symmetry. For example, a plane wave has the following displacement and stress patterns at $r=R$:

$$u_r(\theta) = C_1 i\alpha\cos\theta e^{i\alpha R\cos\theta} \tag{12}$$

$$u_\theta(\theta) = -C_1 i\alpha\sin\theta e^{i\alpha R\cos\theta} \tag{13}$$

$$\sigma_{rr}(\theta) = -C_1(\lambda+2\mu\cos^2\theta)\alpha^2 e^{i\alpha R\cos\theta} \tag{14}$$

$$\sigma_{r\theta}(\theta) = 2C_1\mu\sin\theta\cos\theta\alpha^2 e^{i\alpha R\cos\theta} \tag{15}$$

where C_1 is the given strength of the incident wave. A point source situated on the inclusion's axis of symmetry emitting spherically symmetric waves qualifies to be treated by the current method. This source would have its own displacement and stress representations at $r=R$. In general, the incident wave may be composed of any system of wave harmonics with known generalized coordinates C_j that satisfies the wave equations.

Solution Method

Let m be the number of nodes on the mesh boundary. Let $\mathbf{U}_b^I$ and $\mathbf{U}_b^S$ denote two $(2m\times 1)$ arrays of the radial and axial boundary nodal displacements due to the incident and scattered waves, respectively. Consider that the incident wave is represented by p global functions and assume that the first q global functions for both compressional and shear waves be adequate for capturing the scattered field.

The method is based on two ways for expressing $\mathbf{U}_b^I$ and $\mathbf{U}_b^S$. One way is to evaluate them from the global functions. For the scattered field, $\mathbf{U}_b^S$ may be written as:

$$\mathbf{U}_{bg}^S = \mathbf{G}\,\mathbf{X} \tag{16}$$

where $\mathbf{X}$ is the $(2q\times 1)$ array of unknown generalized coordinates A_n, B_n, i.e.,

$$\mathbf{X}^T = [A_0, B_0, A_1, B_1, \ldots\ldots A_{q-1}, B_{q-1}] \tag{17}$$

and $\mathbf{G}$ is a $(2m\times 2q)$ matrix whose columns are the spatial displacement distributions found by evaluating equations (5) and (6) at the boundary nodes. The subscript g is $\mathbf{U}_{bg}^S$ and is used to identify the method of constructing this displacement vector from the global functions. For the incident field, the representation has the form:

$$\mathbf{U}_{bg}^I = \mathbf{H}\,\mathbf{C} \tag{18}$$

where $\mathbf{C}$ contains the p given amplitudes C_j and $\mathbf{H}$ is a $(2m\times p)$ matrix of the incident wave spatial displacement distributions. Combining equations (16) and (18) gives the boundary nodal displacements derived from the global functions of the exterior field:

$$\mathbf{U}_{bg} = \mathbf{U}_{bg}^I + \mathbf{U}_{bg}^S = \mathbf{H}\,\mathbf{C} + \mathbf{G}\,\mathbf{X} \tag{19}$$

The other way of representing the nodal displacements is to solve the interior problem with the forcing function given by the surface tractions associated with the exterior field. These global function tractions, such as those given by equations (7), (8), (14), and (15) can be converted into consistent load vectors, whose forms for the incident and scattered fields are, respectively:

$$\mathbf{F}_b^I = \mathbf{H}^L\mathbf{C} \tag{20}$$

$$\mathbf{F}_b^S = \mathbf{G}^L\mathbf{X} \tag{21}$$

where the columns of $\mathbf{H}^L$ and $\mathbf{G}^L$ are the nodal force distributions. Inversion of equation (4) with the forcing functions given by equations (20) and (21) yields all the finite element nodal displacements, those of which on the mesh boundary, may be written as:

$$\mathbf{U}_{bf}^I = \mathbf{V}\,\mathbf{C} \tag{22}$$

$$\mathbf{U}_{bf}^S = \mathbf{W}\,\mathbf{X} \tag{23}$$

where the columns of $\mathbf{V}$ and $\mathbf{W}$ are the solutions for the force distributions given by the corresponding columns in $\mathbf{H}^L$ and $\mathbf{G}^L$. The subscript f is used to denote solutions of the interior problem. Combining equations (22) and (23) gives the total mesh boundary displacements of the interior problem:

$$\mathbf{U}_{bf} = \mathbf{U}_{bf}^I + \mathbf{U}_{bf}^S = \mathbf{V}\,\mathbf{C} + \mathbf{W}\,\mathbf{X} \tag{24}$$

Recapitulating at this point, the boundary displacements for both the interior problem and exterior field are available. The incident wave amplitudes are completely given, but those of the scattered wave components remain to be determined. Traction reciprocity at the mesh boundary is satisfied term for term by virtue of the equivalence between the tractions of the exterior field and the consistent load vectors of the interior problem. Only displacement continuity at the mesh boundary remains to be satisfied.

To enforce displacement continuity, set equation (19) equal to equation (24) and rearrange this result so that the unknown amplitude $\mathbf{X}$ appears on the left-hand side and the knowns occur on the right-hand side:

$$\mathbf{A}\,\mathbf{X} = \mathbf{B} \tag{25}$$

where

$$\mathbf{A} = [\mathbf{G} - \mathbf{W}], \qquad \mathbf{B} = [\mathbf{V} - \mathbf{H}]\mathbf{C} \tag{26}$$

Matrix $\mathbf{A}$ is of order $(2m\times 2q)$ and the solution to equation (25) may be obtained by least square error minimization. Define an error ϵ as

$$\epsilon = \mathbf{A}\,\mathbf{X} - \mathbf{B} \tag{27}$$

The square of ϵ is constructed by premultiplication by its transpose:

$$\epsilon^2 = \epsilon^T\epsilon = \mathbf{X}^T\mathbf{A}^T\mathbf{A}\,\mathbf{X} - \mathbf{B}^T\mathbf{A}\,\mathbf{X} - \mathbf{X}^T\mathbf{A}^T\mathbf{B} + \mathbf{B}^T\mathbf{B} \tag{28}$$

Minimizing ϵ^2 with respect to $\mathbf{X}$ yields:

$$[\mathbf{A}^T\mathbf{A}]\mathbf{X} = \mathbf{A}^T\mathbf{B} \tag{29}$$

Solving for $\mathbf{X}$ gives:

$$\mathbf{X} = [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{B} \tag{30}$$

For $2m \le 2q$, the matrix operation $[\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}$ is the generalized inverse of $\mathbf{A}$ and uniqueness of this solution is assured (see reference [14]).

Once $\mathbf{X}$ is determined, the complete field is known. The exterior scattered field is represented by the global functions with $\mathbf{X}$ as the generalized coordinates. The refracted field in the finite element model is related to the solution of the interior problem and $\mathbf{X}$. Consequently, stresses and other physically meaningful quantities may be calculated accordingly.

Examples

As illustrations of the present method, spheroidal and right-circular cylindrical inclusions were considered. The incident wave consisted of a plane compressional wave propagating along the axis of symmetry of these objects. But before discussing these examples, some remarks concerning the effectiveness and accuracy of this method need to be mentioned. A computer code was developed and the problem of a homogeneous, isotropic spherical inclusion was analyzed, for which analytical solutions [15] are available for comparison. To assess the overall effectiveness and explore some of the modeling nuances to acquire confidence in the computer code, a range of parameters was considered, viz. (*i*) relative material properties; (*ii*) relative wave length; and (*iii*) various mesh layouts. Results of the study are reported in [16]. Performance of the method in the interior was gauged by calculating percentage errors of typical displacements and stresses within and on the inclusion. In general, these results

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

were highly satisfactory.[3] For reasonable variation of these parameters, the dominant global coefficients were accurate through four significant figures. This demonstrates the ability of the solution method to capture the far field behavior.

One reason for considering spheroidal inclusions is the availability of results (albeit graphical) in the literature due to Varadan and Varadan [7]. They employed the transition matrix and considered a host medium of titanium and a tungsten carbide inclusion, whose properties are, respectively:

$$\lambda = 96.554, \quad \mu = 40.600, \quad \rho = 4.4222$$
$$\lambda' = 175.23, \quad \mu' = 219.00, \quad \rho' = 13.825 \qquad (31)$$
$$(\lambda, \mu \text{ in GPA and } \rho \text{ in Mg/m}^3)$$

Numerical work was carried for spheroids with aspect ratio $b/a = 0.5, 1, 2$ where a denotes the equatorial radius. Also, the cases of right-circular cylindrical inclusions that circumscribe these spheroids were considered. To the best of the authors' knowledge, no data exist for this type of scatterer. The finite element mesh layouts are shown in Fig. 2. In all cases 25-node isoparametric elements were used. Although singularities occur at the corners of the cylindrical inclusions, they are disregarded in this analysis and this omission should not appreciably affect the scattered field. To determine the appropriate number of global functions, the total number employed was successively increased until the first three digits in the computations of the scattering cross sections (see the next paragraph) showed no change in two consecutive computer runs. The number N of global functions used are shown in Fig. 2.

A quantity of physical interest is the differential scattering cross section, which measures the energy scattered into different directions. Summarized here are the formulas needed to evaluate differential scattering cross section. The reader is referred to Gubernatis et al. [17] for a more complete discussion of the derivation of these formulas. It is known that the axisymmetric scattered field has the following asymptotic representation for the displacements:

$$u_r(r,\theta) = F(\theta)\frac{e^{i\alpha r}}{r} + 0\,(r^{-2})$$
$$u_\theta(r,\theta) = G(\theta)\frac{e^{i\beta r}}{r} + 0\,(r^{-2}) \text{ as } r \to \infty \qquad (32)$$

where $F(\theta)$ and $G(\theta)$ are readily obtained from the global functions. For a purely compressional incident wave, the differential scattering cross section is given by the formula

$$\sigma(\theta) = \sigma^P(\theta) + \frac{\alpha}{\beta}\sigma^S(\theta) \qquad (33)$$

where σ^P and σ^S are the longitudinal and transverse components defined by:

$$\sigma^P(\theta) = F(\theta)F^*(\theta), \quad \sigma^S(\theta) = G(\theta)G^*(\theta) \qquad (34)$$

in which the asterisk denotes complex conjugation.

Polar plots of the total differential scattering cross sections for an incident wave with wave length $L = \pi a$ for both spheroids and cylinders are shown in Figs. 3(a–c). The generally larger values for the cylinders may be attributed to their more blunt nature and greater volume in comparison to their more streamlined spheroidal counterpart. The total differential cross sections for the spheroid and cylinder with $b/a = 0.5$ are fairly similar in pattern and magnitude. In contrast, their patterns and magnitudes differ markedly as the aspect ratio b/a increases.

The values of the longitudinal and transverse components of the differential cross sections, plotted as functions of αa,

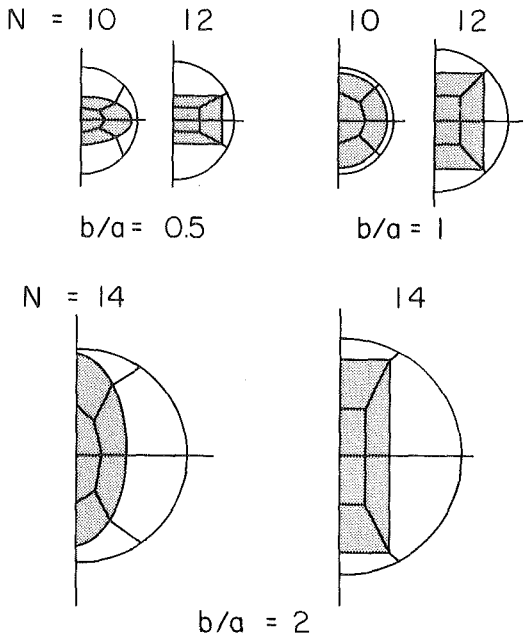


Fig. 2 Mesh layouts for spheroids (left side) and cylinders (right side) with aspect ratios $b/a = 0.5, 1, 2$

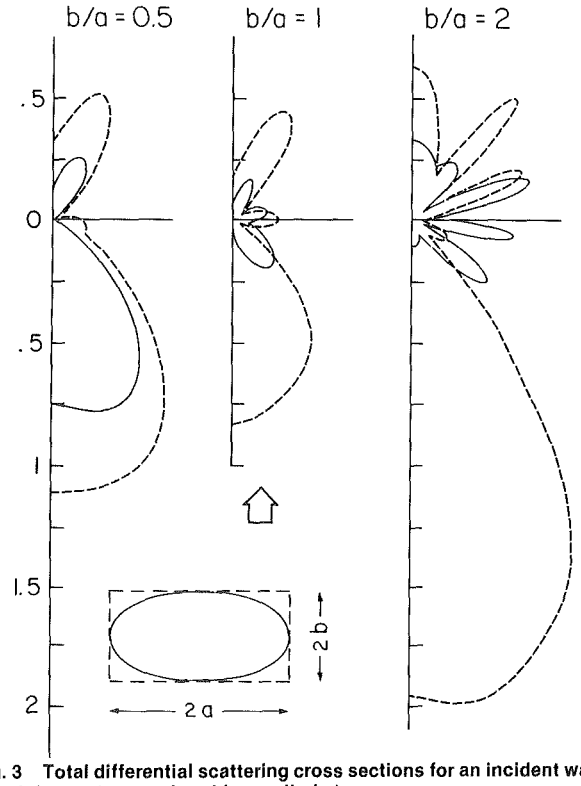


Fig. 3 Total differential scattering cross sections for an incident wave length $L/a = \pi$(——: spheroid, ···: cylinder)

are shown in Fig. 4 for various azimuthal angles θ. A visual comparison of the curves in Fig. 4 for the spheroid with those of Varadan and Varadan (Figs. 3 and 4 of [7]) reveals good agreement, particularly in the long wave length regime.[4]

Concluding Remarks

A method based on global local finite element concepts for

[3] Some problems were experienced in stress calculations which subsequent to completion of [16] were corrected by simple modification of nodal configurations and by introduction of stress extrapolation technique.

[4] A factor of ½ exists between the P-P curves for the spheroid in Fig. 4 and Fig. 3 of reference [7]. This discrepancy was resolved by communication with the authors of [7] and the current results are correct as shown.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

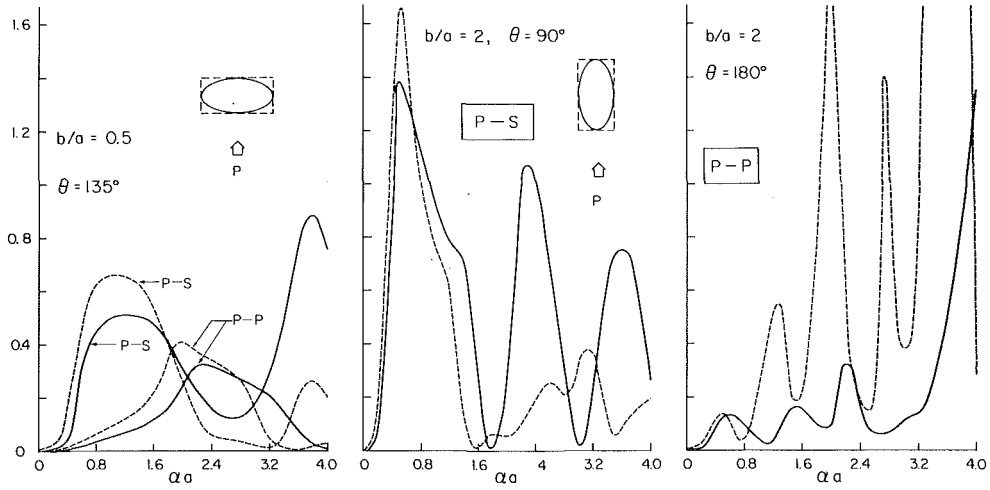


Fig. 4 Longitudinal and transverse differential cross sections (——: spheroid, ---: cylinder)

analysis of axisymmetric elastic scattering by arbitrary axisymmetric objects has been presented. The specific case of steady plane compressional waves illuminating a single object was discussed. Comparison of the present results with available solutions confirms the effectiveness of this method. With the present formulation, more general cases may be treated, for example, inhomogeneous scatters and arbitrary incident compressional waves such as a point source situated on the inclusion's axis of symmetry. Also, the problem of multiple axisymmetric objects, all aligned along the propagation direction, may be treated.

As described, the method calls for satisfying traction continuity at the finite element mesh boundary first and then enforcing displacement continuity by a least square fit. It is mentioned that the reverse procedure is possible. Using equation (4), with $\mathbf{U}_b$ prescribed, $\mathbf{U}_i$ can be determined from which the surface stresses on the finite element mesh boundary can be computed. Then, traction continuity is enforced on a least square basis. This alternate procedure has been tried with equal success, however, the programming details are somewhat more cumbersome.

Mention is made of a similar method used by Morgan and Mei in electromagnetic wave scattering [18][5]. Resemblances include combined use of finite elements and spherical wave functions and a least square error enforcement of continuity. Their method is different in that a triregional subdivision is involved: (1) an inner sphere within the scatterer where diffracted waves are described by spherical wave functions, (2) an intermediate region consisting of the remainder of the scatterer and part of the host medium modeled by finite elements, and (3) the exterior region where outgoing waves are described by spherical wave functions. Inter-regional continuity by least square error is enforced on two interfaces. Therefore, more nodes and more equations are involved. Moreover, their method is restricted to a homogeneous inner core for the scatterer. The present method is more versatile because inhomogeneous, anisotropic scatters can be treated.

In closing, it is noted that this method can be easily extended to incident shear waves as well as to incident waves arriving at oblique angles to the axisymmetric scatterers. The essence of the method is conceptually valid for general three-dimensional objects. The details and the computational feasibilities, however, remain to be explored before any remarks of substance can be made regarding the suitability of such extensions.

[5]This paper was called to the authors' attention by a reviewer, for which sincere appreciation is expressed.

References

1 Pao, Y. H., and Mow, C. C., *Diffraction of Elastic Waves and Dynamic Stress Concentrations*, Crane and Russak, New York, 1973.

2 Waterman, P. C., "Matrix Theory of Elastic Wave Scattering. II. A New Conservation Law," *Journal of the Acoustical Society of America*, Vol. 63, 1978, pp. 1320–1335.

3 Waterman, P. C., "Matrix Theory of Elastic Wave Scattering," *Journal of the Acoustical Society of America*, Vol. 60, 1976, pp. 567–580.

4 Varatharajulu, V., and Pao, Y. H., "Scattering Matrix for Elastic Waves. I. Theory," *Journal of the Acoustical Society of America*, Vol. 60, 1976, pp. 556–566.

5 Pao, Y. H., "The Transition Matrix for the Scattering of Acoustic Waves and for Elastic Waves," *Proceedings of the IUTAM*, Wiley-Interscience, New York, 1978, pp. 123–144.

6 Pao, Y. H., "Betti's Identity and Transition Matrix for Elastic Waves," *Journal of the Acoustical Society of America*, Vol. 64, 1978, pp. 302–310.

7 Varadan, V. V., and Varadan, V. K., "Scattering Matrix for Elastic Waves. III. Application to Spheroids," *Journal of the Acoustical Society of America*, Vol. 65, 1979, pp. 896–905.

8 Visscher, W. M., "A New Way to Calculate Scattering of Acoustic and Elastic Waves I. Theory Illustrated for Scaler Waves," *Journal of Applied Physics*, Vol. 51, 1980, pp. 825–834.

9 Visscher, W. M., "A New Way to Calculate Scattering of Acoustic and Elastic Waves II. Application to Elastic Waves Scattered from Voids and Fixed Rigid Obstacles," *Journal of Applied Physics*, Vol. 51, 1980, pp. 835–845.

10 Muki, R., and Dong, S. B., "Some Remarks on the Use of Asymptotic Solutions in Global Local Finite Element Analysis for an Elastic Half-Space," *Recent Research on Mechanical Behavior of Solids*, University of Tokyo Press, 1979, pp. 55–78.

11 Muki, R., and Dong, S. B., "Elastostatic Far Field Behavior in a Layered Half-Space Under Surface Pressure," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 504–512.

12 Avanessian, V., "Global-Local Finite Element Analysis of Pile-Soil Interaction," MS Thesis, University of California, School of Engineering and Applied Science, Los Angeles, Calif., 1980.

13 Dong, S. B., "Global-Local Finite Element Methods," *State-of-the-Art Survey of Finite Element Methods*, Pilkey, W. D., and Noor, A. K., eds., ASME Publication, Chapter 14, 1981.

14 Lancaster, P., *Theory of Matrices*, Academic Press, New York, 1969, pp. 301–307.

15 Pao, Y. H., and Mow, C. C., "Scattering of Plane Compressional Waves by a Spherical Obstacle," *Journal of Applied Physics*, Vol. 34, 1963, pp. 493–499.

16 Goetschel, D. B., "Analysis of Steady State Elastic Wave Scattering by Global-Local Finite Elements," Ph.D Dissertation, University of California, Los Angeles, Calif., 1980.

17 Gubernatis, J. E., Donany, E., and Krumahansl, J. A., "Formal Aspects of the Theory of the Scattering of Ultrasound by Elastic Materials," *Journal of Applied Physics*, Vol. 48, 1977, pp. 2804–2811.

18 Morgan, M. A., and Mei, K. K., "Finite-Element Computation of Scattering by Inhomogeneous Penetrable Bodies of Revolution," *IEEE Transactions on Antennas and Propagation*, Vol. AP-27, No. 2, Mar. 1979, pp. 202–214.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

R. L. Weaver[1]

Yih-Hsing Pao
Fellow ASME

Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, N.Y. 14853

Axisymmetric Elastic Waves Excited by a Point Source in a Plate

The response of an infinite elastic plate to dynamic loading is presented by the method of superposition of normal modes, a method particularly appropriate in the intermediate and far field. The method is compared with the method of integral transforms. Explicit expressions are given for the case of loading by a concentrated vertical step force. These expressions are evaluated numerically over a range of distances from 4 to 40 plate thicknesses. The numerical results are compared with qualitative stationary phase analyses and with the exact results of generalized ray theory.

1 Introduction

The propagation of waves in an infinite elastic plate has been studied since 1889 when Rayleigh [1] and Lamb [2] established the frequency equation for the dispersive waves based on the theory of elasticity. Since then there have been numerous contributions, notably those of Mindlin [3] on the dispersion relations and approximate theories and Miklowitz [4] on transient waves.

Most of the transient solutions are derived by applying double or triple integral transforms. The inverse transformation integrals are then evaluated either by the calculus of residues plus the method of stationary phase [5–7], or by the method of generalized rays [8–12]. As can be seen from the text by Miklowitz [4], the evaluation of these integrals in the first approach involves the understanding of the intricate behavior of dispersion relations for real as well as complex wave numbers. So far very little quantitative information about the wave motion at intermediate range (10–100 times the thickness of the plate from the source) or far range (100 times the thickness or longer) are available.

As for the application of the method of generalized rays, earlier results [8–10] were confined to epicentral response or wave motion very close to the source and epicenter. Only recently numerical results were reported for wave motion in near range (0 to 10 thickness from the source) [11, 12].

The theory of generalized rays is an exact method, as distinguished from the geometric ray theories. It becomes extremely cumbersome, though, for a receiver-source separation of greater than 10 plate thicknesses because a very large number of generalized ray integrals must be evaluated numerically. The paper by Ceranoglu and Pao [12] showed that 1586 ray integrals were calculated and summed at a range of six plate thicknesses over a duration of 10 transit times (one transit time is the time for the dilatational wave to cross the plate thickness once). Although all these ray integrals are assembled and calculated automatically by using one large computer code, it is not very practical when long time and far field results of wave motion are desired.

In this paper, we present another approach, the method of expansion in eigenfunctions, known also as the method of expansion in normal modes. This method, usually applied to finite bodies, is discussed in the Treatise (Sections 124-128) by Love [13]. A comprehensive treatment of the method can be found in the text (Section 5.17) by Eringen and Suhubi [14] and an application to finite bodies in the work of Reismann [15]. The method gives precisely the same expressions as can be obtained by integral transforms. The resulting expressions, as shown by numerical results in this paper, are particularly suitable for calculating intermediate and long range wave fields.

Since the concept of normal modes is well defined for a body bounded in all directions, we determine the normal modes of a thick, circular plate with a finite radius in free vibration in Sections 2 and 3, and formulate the solution of forced vibration of the plate by the method of normal mode expansion in Section 4. By letting the radius of the plate approach infinity, we then obtain, in Section 5, the solution for the transient waves in an infinite plate. Evaluation of the resulting integral expressions by the method of stationary phase is discussed in Section 7. Numerical evaluations are presented in Section 8. Throughout this paper, we consider only axially symmetric waves.

As discussed in Section 6, the method of normal mode expansion is easily applicable to surface forces as well as buried forces in a plate. Our interest in a solution for buried sources stems from developments in acoustic emission. With the advent of increasing sophistication in acoustic emission techniques, theoretical results for the dynamic response of a structure such as a large plate over a wide range of times and positions are needed to locate and to characterize the buried

[1] Present address: Department of Theoretical and Applied Mechanics, 216 Talbot Laboratory, 104 South Wright Street, University of Illinois, Urbana, Ill. 61801.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May, 1981; final revision, March, 1982. Paper No. 82-WA/APM-15.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

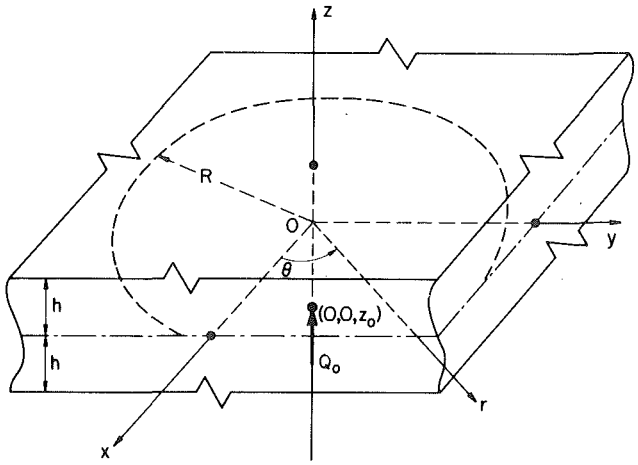


Fig. 1 Axisymmetric force in a thick plate

sources of emission. Of particular interest is the identification of those aspects of measured responses that are affected by the geometry of the plate, the source characteristics, or the receiver characteristics. This identification cannot be accomplished without a detailed theoretical analysis and numerical results.

2 Basic Equations

We consider an infinite elastic plate bounded by two parallel surfaces free of tractions at $z = \pm h$ (Fig. 1), $z = 0$ being the midplane of the plate. The circular cylindrical coordinates r, θ, z will be used to describe the response of the plate to an axisymmetric time-dependent force distribution.

The linear elastic wave equation is

$$\mathbf{L}(\mathbf{u}) \equiv (\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}) + \mu\nabla^2\mathbf{u} = \rho\ddot{\mathbf{u}} - \mathbf{F}(\mathbf{x},t) \tag{2.1}$$

where $\mathbf{u}(\mathbf{x},t)$ is the time-dependent displacement at position $\mathbf{x}$, and $\mathbf{F}(\mathbf{x},t)$ is an externally applied body force density, usually taken equal to zero for $t \leq 0$. ρ is the mass density, and λ and μ the Lamé constants of the isotropic plate material, ρ, λ, and μ being assumed independent of $\mathbf{x}$ and t. The operator $\mathbf{L}$ is defined as $\mathbf{L} \equiv (\lambda+\mu)\nabla\nabla\cdot + \mu\nabla^2$. In the ensuing discussions we will consider only axisymmetric disturbances where all field quantities are independent of angular coordinate θ, and the θ component of displacement is zero.

The boundary conditions at the two parallel surfaces are

$$\sigma_{zr} = \sigma_{zz} = 0 \quad \text{at} \quad z = \pm h \tag{2.2}$$

In addition to these boundary conditions, there are also the initial conditions

$$\mathbf{u}(\mathbf{x},0) = \dot{\mathbf{u}}(\mathbf{x},0) = 0 \tag{2.3}$$

In many cases where the applied force is confined to the surfaces at $z = \pm h$, one may consider the equivalent problem obtained by setting $\mathbf{F}$ to zero but employing inhomogeneous time-dependent boundary conditions for the stresses at $z = \pm h$. As shown later, the case of surface loading can be treated as a special case of a general class of plate problems formulated by equations (2.1), (2.2), and (2.3).

Before considering the case of forced motion, we discuss first the free vibration of an elastic plate, $\mathbf{F} = 0$ in equation (2.1). We will investigate the simple harmonic motion of the plate, that is, solutions with time dependence $\exp(-i\omega t)$ for some real ω. This time-dependent factor will be suppressed throughout the following.

The problem of plane harmonic waves [1, 2] and circular-crest harmonic waves [16] in an infinite plate has been exhaustively examined [3, 17, 18]. We shall, however, repeat some of the results which are needed for ensuing discussions.

In addition to the boundary conditions at $z = \pm h$, we have the condition that $\mathbf{u}$ be regular at $\mathbf{r} = 0$. The boundary conditions at large r will be considered in Section 3. The harmonic solutions to equation (2.1) are then, for any κ

(1) Antisymmetric modes

$$\begin{aligned} u_r &= (B\sin\alpha z - C\beta\sin\beta z)\,\kappa J_0'(\kappa r) \\ u_z &= (B\alpha\cos\alpha z + C\kappa^2\cos\beta z)\,J_0(\kappa r), \end{aligned} \tag{2.4}$$

where

$$B/C = 2\kappa^2\beta\sin\beta h/(\kappa^2 - \beta^2)\sin\alpha h, \tag{2.5}$$

and ω satisfies

$$R_a(\omega,\kappa) \equiv (\kappa^2 - \beta^2)^2\sin\alpha h\cos\beta h + 4\alpha\beta\kappa^2\cos\alpha h\sin\beta h = 0. \tag{2.6a}$$

(2) Symmetric modes

$$\begin{aligned} u_r &= (A\cos\alpha z + D\beta\cos\beta z)\,\kappa J_0'(\kappa r) \\ u_z &= (-\alpha A\sin\alpha z + D\kappa^2\sin\beta z)\,J_0(\kappa r) \end{aligned} \tag{2.7}$$

where

$$A/D = -2\kappa^2\beta\cos\beta h/(\kappa^2 - \beta^2)\cos\alpha h \tag{2.8}$$

and ω satisfies

$$R_s(\omega,\kappa) \equiv (\kappa^2 - \beta^2)^2\cos\alpha h\sin\beta h + 4\alpha\beta\kappa^2\cos\beta h\sin\alpha h = 0. \tag{2.6b}$$

In the preceding

$$\alpha^2 = \omega^2/c_p^2 - \kappa^2, \quad \beta^2 = \omega^2/c_s^2 - \kappa^2 \tag{2.9}$$

where c_p and c_s are the dilatational (p) and shear (s) wave velocities. The first of the foregoing results describes an antisymmetric motion because the vector displacement field is antisymmetric upon reflection through the midplane $z = 0$. The second describes a symmetric displacement field.

Equations (2.6a,b) are the same as the classical frequency equations for straight-crest waves (plane strain) derived by Rayleigh in 1889 and analyzed by Lamb, giving frequency ω as an implicit transcendental function of κ, or vice versa. Equations (2.4)–(2.9) for real κ determine the mode shapes of free waves in an infinite plate, up to an arbitrary scale factor associated with each mode.

The behavior of the solution of equations (2.6) for complex as well as real values of κ has been extensively investigated [3]. For the present purposes, it will be sufficient to seek solutions for real wave numbers κ. Figure 2 shows the $\omega(\kappa)$ dependence as determined numerically for moderate and small values of real κ for the case of a glass plate of Poisson ratio, ν, equal to 0.21. Note that the dispersion relation is displayed by many branches on the $\omega - \kappa$ plane.

In the following, the branches will be ordered by numerals $m = 1,2,3,4, \ldots$, with odd integers for antisymmetric modes (dashed lines labeled by A_i in Fig. 2) and even integers for symmetric modes (solid lines labeled by S_i). They will also be referred to by their symmetry type and their order as given by the magnitude of their cutoff frequencies. Thus $m = 1$ is for the first antisymmetric branch A_1, and $m = 2$ for the first symmetric branch S_1, both starting at origin. The $m = 5$ designates the third antisymmetric branch, A_3, with κ ranging over real values from zero to infinity and ω/ω_s ranging from 3.0 to infinity, the $\omega_s (= \pi c_s/2h)$ being the frequency of the lowest thickness shear mode [3].

The concept of distinct branches is somewhat artificial in that the branches are continuosly connected from one to another by paths amongst complex values of κ. The existence of a multitude of branches indicates that at given κ, there are an infinite number of modes of wave propagation of the plate at different frequencies. At $\kappa = 0$, $\alpha = \omega/c_p$, and $\beta = \omega/c_s$ the wave normal is along the x-axis and the modes are called simple thickness modes [3].

Figures 3(a) and 3(b) show the radial group velocities $c_g = \partial\omega/\partial\kappa$ implied by the dispersion curves of Fig. 2. Note that the

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

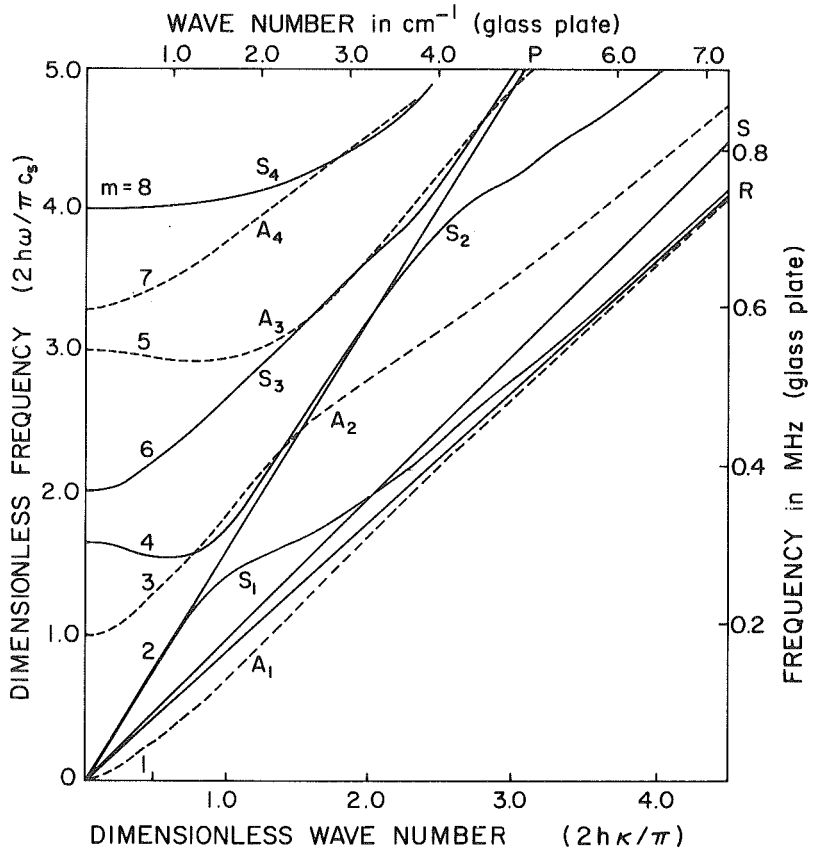


Fig. 2 First eight branches of dispersion relation for Rayleigh-Lamb waves in a plate ($\nu = 0.21$). (Glass plate: $2h = 0.960$ cm, $c_p = 0.576$ cm/μsec, $c_s = 0.349$ cm/μsec.) Erratum: the scale of the horizontal axis on the top should be doubled, and the wave number in cm^{-1} from 1.0 to 7.0 should be changed to 2.0 to 14.0.

Table 1 Maximum and minimum group velocities ($\nu = 0.21$) ($\omega_s = \pi c_s/2h$)

	Maxima		Minima	
Branch	c_g/c_s	ω/ω_s	c_g/c_s	ω/ω_s
A_1	0.99	0.93	–	–
A_2	1.17	1.80	0.68	3.01
A_3	0.14	3.04	0.71	5.73
	1.26	4.12		
A_4	0.79	4.22	0.77	4.69
S_1	1.59	0.00	0.49	1.55
S_2	0.21	1.61	0.70	4.38
	1.55	2.63		
S_3	0.97	2.95	0.93	3.53
	1.55	5.01	~0.6	~7.0

first symmetric branch (S_1 in Fig. 3(b) and first antisymmetric branch (A_1 in Fig. 3(a)) have asymptotic group velocities equal to the Rayleigh surface wave speed in a half space (slope of line OR in Fig. 2). Thus the modes lying on these branches at large κ are essentially Rayleigh surface waves. All other branches have asymptotic group velocities equal to the shear wave speed (slope of line OS in Fig. 2). At no point does any group velocity rise above the dilitational wave speed (slope of line OP in Fig. 2).

The maximum and minimum values of group velocities for each branch are listed in Table 1. In addition to the points indicated here, all branches have asymptotic stationary values of group velocity at infinite frequency, at, for S_1 and A_1, the Rayleigh wave speed $c_g/c_s = 0.9127$ and for other branches at the shear wave speed $c_g/c_s = 1.000$. The physical significance of these values will be discussed in Section 7.

3 Normal Modes of a Large Circular Plate

The expressions in equations (2.4) and (2.7) are the modes of free wave propagation in an infinite plate. When the plate is bounded by a cylindrical surface $r = R$, additional restrictions on the wave number κ will be imposed by the boundary conditions at $r = R$. The familiar boundary conditions of a traction-free surface ($\sigma_{rr} = \sigma_{rz} = 0$) or a rigid surface ($u_r = u_z = 0$), however, cannot be satisfied by these solutions [3, 19]. On the other hand, these solutions can be made to satisfy a rigid-smooth boundary condition, that is,

$$u_r = \sigma_{rz} = 0 \text{ at } r = R \tag{3.1}$$

The condition (3.1) can be shown to restrict κ to a set of discrete real irrational numbers,

$$\kappa = \kappa_n \equiv \gamma_n/R, \quad n = 1,2,\ldots \tag{3.2}$$

where γ_n are the roots of the radial frequency equation

$$J_0'(\gamma_n) = -J_1(\gamma_n) = 0. \tag{3.3}$$

For large n, the roots κ_n are nearly uniformly spaced and

$$\Delta\kappa = \kappa_{n+1} - \kappa_n \cong \pi/R \tag{3.4}$$

Hence, as we shall see later, the spacings of the roots are infinitesimally close when R approaches infinity.

From Fig. 2 we notice that for each value of $\kappa = \kappa_n$, there are an infinite number of values of ω, designated ω_{mn}. The

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

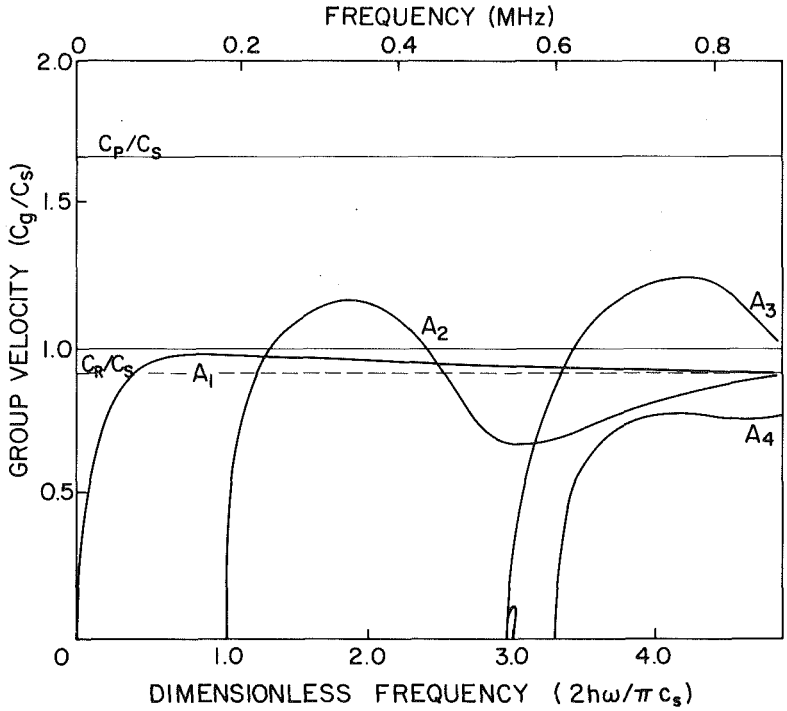


Fig. 3(*a*) Group velocity curves for antisymmetric branches

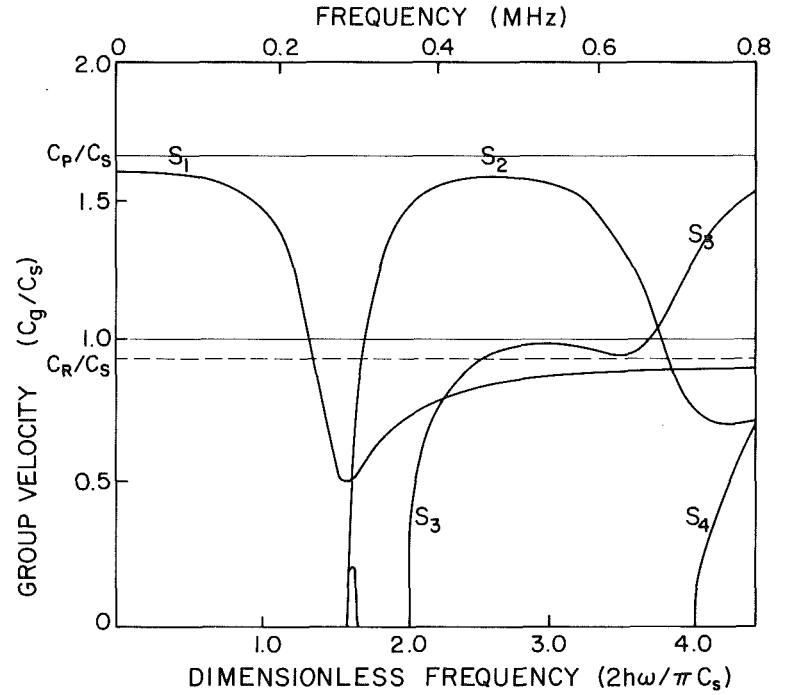


Fig. 3(*b*) Group velocity curves for symmetric branches

Fig. 3

index m corresponds to the order of branches and the ω_{mn} are the roots of Rayleigh-Lamb equation

$$R_{a,s}(\omega_{mn}, \kappa_n) = 0 \quad m, n = 1, 2, 3, \ldots \tag{3.5}$$

When α and β are calculated from equation (2.9) for each κ_n and ω_{mn}, the displacement vector of equation (2.4) represents a principal antisymmetric mode of vibration of a thick plate with finite radius R, and that of equation (2.7) a principal symmetric mode.

Since the two amplitudes B, C in equation (2.4) and A, D in equation (2.7) bear a constant ratio, we choose, according to equations (2.5) and (2.8)

$$B = 2\kappa^2 \beta \sin \beta h$$

$$C = (\kappa^2 - \beta^2) \sin \alpha h \tag{3.6}$$

for the antisymmetric modes, and

$$A = -2\kappa^2 \beta \cos \beta h$$

$$D = (\kappa^2 - \beta^2) \cos \alpha h \tag{3.7}$$

for the symmetric modes. When all κ's are appended with the subscript n, and all α's, β's, and ω's are appended with double subscripts mn, the principal modes are expressed by

$$\mathbf{u}_{mn}(\mathbf{x}) = U_{mn}(\mathbf{x})\hat{\mathbf{r}} + W_{mn}(\mathbf{x})\hat{\mathbf{z}} \tag{3.8}$$

In the foregoing, we have for antisymmetric modes (m = 1, 3, 5, . . .)

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

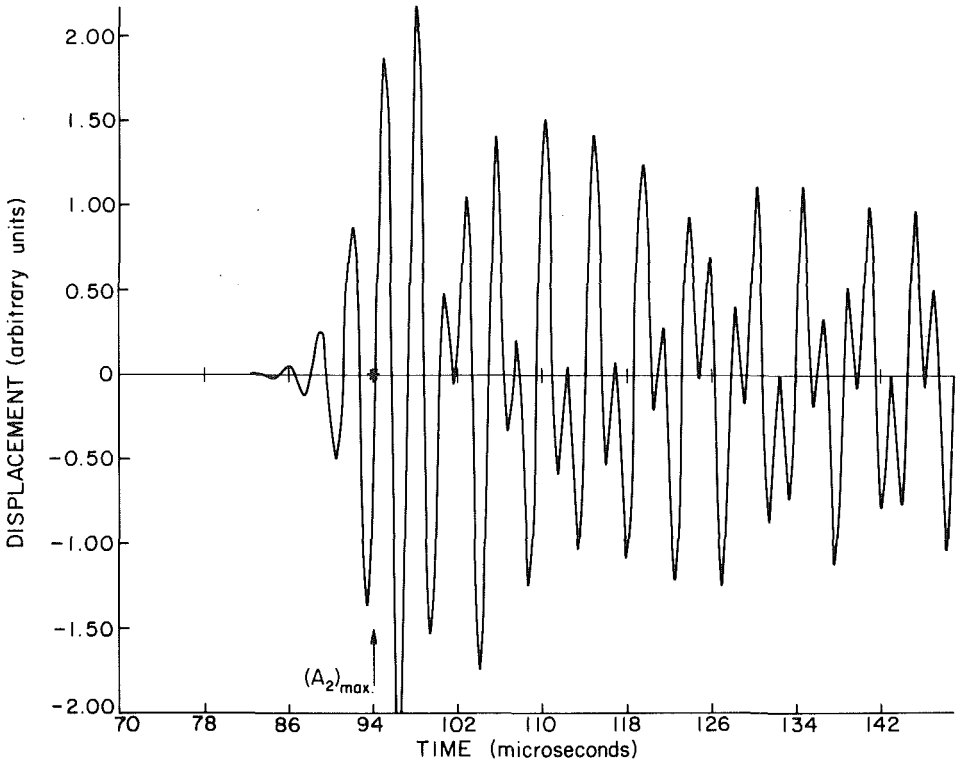


Fig. 4(*a*) The Airy phase of the second antisymmetric branch with the maximum group velocity $c_g = 1.17\, c_s$, $(A_2)_{max}$

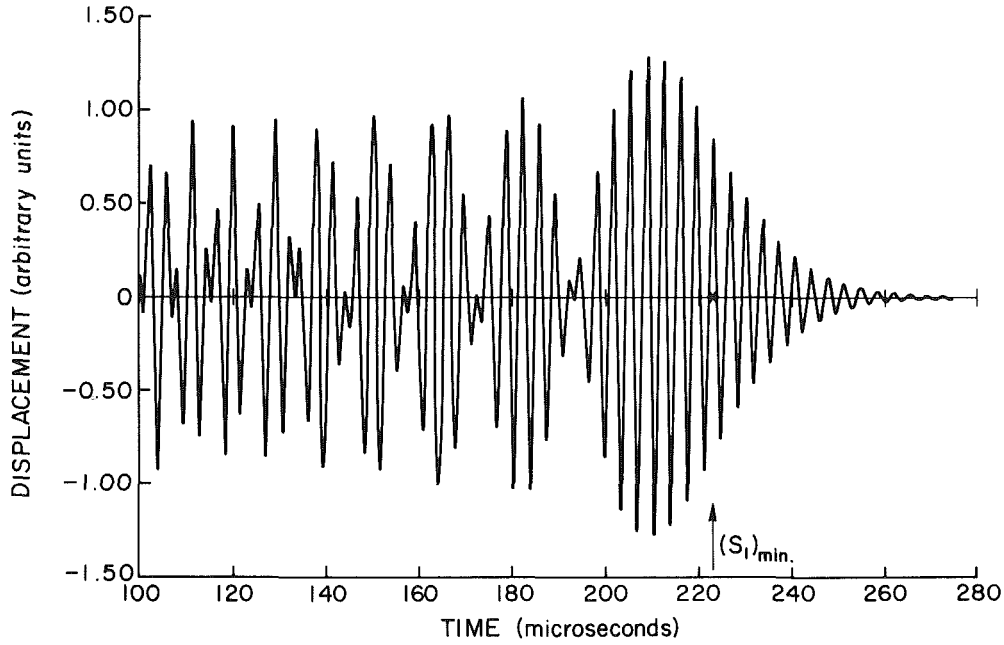


Fig. 4(*b*) The Airy phase of the first symmetric branch with the minimum group velocity $c_g = 0.49\, c_s$, $(S_1)_{min}$

Fig. 4 The Airy phase of the normal displacement at $r = 80\, h$.

$$U^a_{mn} = \beta_{mn}[2\kappa_n^2 \sin \beta_{mn}h \sin \alpha_{mn}z - (\kappa_n^2 - \beta_{mn}^2)\sin \alpha_{mn}h \sin \beta_{mn}z]\kappa_n J_0'(\kappa_n r)$$

$$W^a_{mn} = \kappa_n^2[2\alpha_{mn}\beta_{mn} \sin \beta_{mn}h \cos \alpha_{mn}z + (\kappa_n^2 - \beta_{mn}^2)\sin \alpha_{mn}h \cos \beta_{mn}z]J_0(\kappa_n r), \quad (3.9)$$

and for symmetric modes ($m = 2, 4, 6, \ldots$),

$$U^s_{mn} = \beta_{mn}[-2\kappa_n^2 \cos \beta_{mn}h \cos \alpha_{mn}z + (\kappa_n^2 - \beta_{mn}^2)\cos \alpha_{mn}h \cos \beta_{mn}z]\kappa_n J_0'(\kappa_n r)$$

$$W^s_{mn} = \kappa_n^2[2\alpha_{mn}\beta_{mn} \cos \beta_{mn}h \sin \alpha_{mn}z + (\kappa_n^2 - \beta_{mn}^2)\cos \alpha_{mn}h \sin \beta_{mn}z]J_0(\kappa_n r), \quad (3.10)$$

As in equation (2.9), α and β are determined by

$$\alpha_{mn}^2 = \omega_{mn}^2/c_p^2 - \kappa_n^2$$

$$\beta_{mn}^2 = \omega_{mn}^2/c_s^2 - \kappa_n^2. \quad (3.11)$$

When properly normalized by a constant, $\mathbf{u}_{mn}(\mathbf{x})$ are also called the *normal modes* of the vibrating plate.

Restoring the time factor $\exp(-i\omega t)$ and superposing all

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

normal modes, we can express the general motion of the freely vibrating plate,

$$\mathbf{u}(\mathbf{x}, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{mn}[U_{mn}(\mathbf{x})\hat{\mathbf{r}} + W_{mn}(\mathbf{x})\hat{\mathbf{z}}]e^{-i\omega_{mn}t} \quad (3.12)$$

The circular plate is free of stresses on the two flat surfaces (equation 2.2) and is rigidly and smoothly constrained at the circumference (equation 3.1). The complex coefficients, A_{mn}, are to be determined from the hitherto unspecified initial conditions.

4 Forced Vibration of a Large Circular Plate

In this section, we present the formal solution of the motion of a finite plate under the excitation of a time-dependent body force $\mathbf{F}(\mathbf{x}, t)$. The motion is governed by equation (2.1) with boundary conditions (2.2) and (3.1), and initial conditions (2.3). The solution is a special case of a general class of solutions for elastodynamic equations derived by the method of eigenfunction expansions [13–15]. We summarize this method in the following subsection and present the solution for the circular plate in the next one.

4.1 Method of Eigenfunction Expansion. A linear differential operator $\mathbf{L}$ is self-adjoint in a vector space S of vector functions defined in a volume Ω with respect to the inner product

$$(\mathbf{u}, \mathbf{v}) = \int_{\Omega} \mathbf{u} \cdot \mathbf{v} dV \quad (4.1)$$

if for all $\mathbf{u}$ and $\mathbf{v}$ in S

$$\int_{\Omega} \mathbf{v} \cdot \mathbf{L}(\mathbf{u})\, dV = \int_{\Omega} \mathbf{u} \cdot \mathbf{L}(\mathbf{v})\, dV \quad (4.2)$$

Restricting the space S to suitably continuous functions that satisfy homogeneous boundary conditions on the boundary of Ω, one finds that the operator $\mathbf{L}$ defined by equation (2.1) is self-adjoint (and so is its generalization to anisotropic and inhomogeneous media) for the following boundary conditions:

At each point of the boundary, let $\hat{\mathbf{e}}_i$ be three mutually perpendicular directions. The self-adjoint boundary conditions are

$$\hat{\mathbf{e}}_i \cdot [a_i \mathbf{u} + b_i \mathbf{T}] = 0, \quad (i = 1, 2, 3) \quad (4.3a)$$

$\mathbf{T}$ being the surface traction on the boundary associated with the field $\mathbf{u}$. The a_i and b_i may take any values as long as they are not both zero at the same point of the boundary. Traction-free boundary conditions, for example, may be represented by $a_i = 0$, $b_i = 1$ for all i. Rigid-smooth boundary conditions may be represented by $a_i = 1$, $b_i = 0$ for the rigid normal direction $\hat{\mathbf{e}}_i = \mathbf{n}$, and by $a_j = 0$, $b_j = 1$ for the smoothly sliding directions $\hat{\mathbf{e}}_j$ = tangent directions. We may express these boundary conditions symbolically as

$$\mathbf{B}(\mathbf{u}, \mathbf{T}) = 0 \quad \text{on boundary.} \quad (4.3b)$$

Let $\mathbf{u} = \mathbf{u}^{(m)}(\mathbf{x}) \exp(-i\omega_m t)$ be the solution of equation (2.1) without the body force $\mathbf{F}$, satisfying the boundary conditions (4.3), that is,

$$\mathbf{L}(\mathbf{u}^{(m)}) = -\rho\omega_m^2 \mathbf{u}^{(m)} \text{ in } \Omega \quad (4.4)$$

and

$$\mathbf{B}(\mathbf{u}^{(m)}, \mathbf{T}^{(m)}) = 0 \text{ on boundary} \quad (4.5)$$

The $\mathbf{u}^{(m)}$ are then the eigenfunctions of the operator $\mathbf{L}$, and ω_m the eigenvalues. For self-adjoint boundary conditions, the eigenfunctions form an orthogonal set with the weighting function ρ.

$$\int_{\Omega} \rho \mathbf{u}^{(m)} \cdot \mathbf{u}^{(n)} dV = 0 \quad m \neq n \quad (4.6)$$

and all eigenvalues ω_m^2 can be shown to be real and non-negative [13]. The norm of the eigenfunctions is given by

$$M^{(m)} = \int_{\Omega} \rho \mathbf{u}^{(m)} \cdot \mathbf{u}^{(m)} dV \quad (4.7)$$

Since the eigenfunctions form a complete set [14], any arbitrary functions in the space S can be expanded in a series of these orthogonal eigenfunctions. This enables one to solve the inhomogeneous equation (2.1) when $\mathbf{F} \neq 0$. Let

$$\mathbf{u}(\mathbf{x}, t) = \sum_{m} \mathbf{u}^{(m)}(\mathbf{x})\, q_m(t) \quad (4.8)$$

where $q_m(t)$ are time-dependent coefficients. Substituting the series into equation (2.1), multiplying (inner product) by $\mathbf{u}^{(n)}(x)$, integrating the product over the entire volume Ω, and invoking the orthogonality conditions (4.6), we find

$$M^{(m)}\left[\frac{d^2 q_m{}^{(t)}}{dt^2} + \omega_m^2(t) q_m(t)\right] = F^{(m)}(t) \quad (4.9)$$

where

$$F^{(m)}(t) = \int_{\Omega} \mathbf{u}^{(m)}(\mathbf{x}) \cdot \mathbf{F}(\mathbf{x}, t) dV \quad (4.10)$$

With quiescent initial conditions equation (2.3), we set

$$q_m(0) = \dot{q}_m(0) = 0. \quad (4.11)$$

The general solution of equation (4.9) is then

$$q_m(t) = \frac{1}{\omega_m M^{(m)}} \int_0^t F^{(m)}(\tau) \sin \omega_m (t - \tau)\, d\tau \quad (4.12)$$

Substitution of the foregoing into equation (4.8) completes the solution of equation (2.1) with boundary conditions (4.3), and initial conditions (2.3).

4.2 Transient Motion of a Circular Plate. The normal modes given in equations (3.8)–(3.10) are for a circular plate bounded by $z = \pm h$ and $r = R$. The boundary conditions are specified by equation (2.2) and equation (3.1). It has already been demonstrated that traction-free and rigid-smooth boundary conditions are self-adjoint. Thus, the boundary conditions for the present circular plate are self-adjoint. According to equations (4.6) and (4.7) the normal modes $\mathbf{u}_{mn}(\mathbf{x})$ of equation (4.8) are orthogonal,

$$2\pi \int_{-h}^{h} \int_0^R [\rho \mathbf{u}_{mn}(\mathbf{x}) \cdot \mathbf{u}_{pq}(\mathbf{x})] r dr dz = \delta_{mp} \delta_{nq} M^{(mn)} \quad (4.13)$$

where δ_{mn} is the Kronecker delta.

The transient motion of the plate that is quiescent at $t = 0$ is derived from equations (4.8), (4.10), (4.12),

$$\mathbf{u}(r, z, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{M^{(mn)}} \mathbf{u}_{mn}(r, z) Q^{(mn)}(t) \quad (4.14)$$

where

$$Q^{(mn)}(t) = \frac{1}{\omega_{mn}} \int_0^t F^{(mn)}(\tau) \sin \omega_{mn}(t - \tau)\, d\tau \quad (4.15)$$

and

$$F^{(mn)}(t) = 2\pi \int_{-h}^{h} \int_0^R \mathbf{F}(r, z, t) \cdot \mathbf{u}_{mn}(r, z) r dr dz \quad (4.16)$$

The body force $\mathbf{F}(r, z, t)$ has not yet been specified but for the present problem it must be axially symmetric; the normal modes $\mathbf{u}_{mn}(r, z)$ are given by equation (3.9) when $m = 1, 3, 5 \ldots$, and by equation (3.10) when $m = 2, 4, 6 \ldots$.

The normalization constant $M^{(mn)}$ can be analytically evaluated in this case,

$$M^{(mn)} = 2\pi \int_{-h}^{h} \int_0^R [\rho \mathbf{u}_{mn}(r, z) \cdot \mathbf{u}_{mn}(r, z)] r dr dz \quad (4.17)$$

The part that depends on r can be evaluated by applying the integral formula [formula 11.4.2, reference 20]

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\int_0^R J_\nu^2(\kappa_n r)\,r dr = \frac{1}{2}R^2\begin{cases}(J_0(\kappa_n R))^2 & \nu=0\\ (J_1'(\kappa_n R))^2 & \nu=1\end{cases} \tag{4.18}$$

where $\kappa_n R$ are the roots of $J_0'(\kappa R) = -J_1(\kappa R) = 0$. The z-dependent parts are common integrals. The final results are presented in the following, for sufficiently large R that $(J_0(\kappa_n R))^2 \approx (J_1'(\kappa_n R))^2 \approx 2/(\pi\kappa_n R)$ and for $M_\infty^{(mn)}$ defined by

$$M^{(mn)} = RM_\infty^{(mn)}. \tag{4.19}$$

For m odd (antisymmetric modes)

$$M_\infty^{(mn)} = \frac{2\rho}{\kappa}[B^2(\omega^2 h/c_p^2 + (\alpha^2-\kappa^2)\sin 2\alpha h/2\alpha) + C^2\kappa^2(\omega^2 h/c_s^2 - (\beta^2-\kappa^2)\sin 2\beta h/2\beta) + 4BC\kappa^2 \sin\alpha h \cos\beta h] \tag{4.20a}$$

and for m even (symmetric modes)

$$M_\infty^{(mn)} = \frac{2\rho}{\kappa}[A^2(\omega^2 h/c_p^2 - (\alpha^2-\kappa^2)\sin 2\alpha h/2\alpha) + D^2\kappa^2(\omega^2 h/c_s^2 + (\beta^2-\kappa^2)\sin 2\beta h/2\beta) + 4AD\kappa^2 \cos\alpha h \sin\beta h]. \tag{4.20b}$$

In the foregoing, κ must be understood as having index n and A, B, C, D, α and β as having indices m and n. A, B, C, and D are defined in equations (3.6) and (3.7).

5 Transient Waves in an Infinite Plate

We return to the original problem as posed in Section 2 of an infinite plate excited by a body force $\mathbf{F}(\mathbf{x}, t)$. The solution can be constructed directly from that of a finite plate, as given in the preceding section by letting the radius R approach infinity. Physically, this means that we are ignoring the reflection of waves from the edge at $r = R$, or considering only the time span before the waves generated by the source reach the reflecting edge. Hence, it is immaterial what type of boundary conditions are imposed at the edge.

For some applications one need not take the limit $R \to \infty$, as long as R is large enough that reflections from the outer boundary do not arrive within the times of interest. If R is taken merely large, but not infinite, one is left with a sum over discrete modes instead of the integral, equation (5.1). The integral is a more convenient form for numerical and analytic study, as one need not solve the transcendental equation (3.3) and because one may employ FFT techniques for its evaluation. The integral form is also more generally applicable to very late times or very large distances, without the restrictions $r < R$, $t < (2R-r)/c_p$. Nevertheless, as the numerical evaluation of integrals is always an approximation, the sum form, with $R \neq \infty$, may be more accurate for numerical work.

5.1 General Solutions. As R approaches infinity, the normalization constant $M^{(mn)}$ that contains a factor R^{-1} becomes infinitesimal. Thus each mode in the sum of equation (4.14) contributes an infinitesimal amount to the total amplitude. But in the same limit, the number of modes indexed by n for each value of m is increasing because the spacing between two successive roots, $\Delta\kappa \cong \pi/R$ in equation (3.4), approaches zero. Therefore, in the limit as $R \to \infty$, the summation over the index n of equation (4.14) is replaced by an integration over the continuous wave number κ. The integration is carried out for each branch, indexed by m, from $\kappa = 0$ to $\kappa = \infty$.

Replacing first $1/M^{(mn)}$ in equation (4.14) by $\Delta\kappa/\pi M_\infty^{(mn)}$, taking the limit $\Delta\kappa \to 0$ while $R \to \infty$, and then dropping all subscripts n, we obtain

$$\mathbf{u}(r, z, t) = \sum_{m=1}^{\infty}\int_0^\infty \frac{\mathbf{u}_m(r, z; \kappa)}{\pi M_\infty^{(m)}(\kappa)} Q^{(m)}(t; \kappa)\,d\kappa \tag{5.1}$$

where

$$Q^{(m)}(t; \kappa) = \frac{1}{\omega_m(\kappa)}\int_0^t F^{(m)}(\tau;\kappa)\sin[\omega_m(\kappa)(t-\tau)]d\tau \tag{5.2}$$

$$F^{(m)}(t; \kappa) = 2\pi\int_{-h}^{h}\int_0^\infty \mathbf{F}(r, z, t)\cdot\mathbf{u}_m(r, z; \kappa)\,r dr dz \tag{5.3}$$

Note that the normal modes $\mathbf{u}_m(r, z; \kappa)$ in the foregoing are the same as those given in equations (3.6)–(3.10) except that the discrete wave numbers κ_n are replaced by a continuous variable κ. For each value of κ, there is still a discrete set of values for ω, and hence also for α and β, each indexed by the subscript m. On the other hand, for each m, ω varies continuously with κ as governed by the Rayleigh-Lamb equations (2.6)

$$R_a(\omega_m, \kappa) = 0 \qquad m = 1, 3, 5, \ldots$$

$$R_s(\omega_m, \kappa) = 0 \qquad m = 2, 4, 6, \ldots \tag{5.4}$$

In the final expressions (5.1)–(5.3), $\mathbf{u}_m(r, z; \kappa)/M_\infty^{(m)}(\kappa)$, where $M_\infty^{(m)}(\kappa)$ is given by equation (4.20), represents the spatial variation of the mode indexed by the branch number, m, and the radial wave number κ. $F^{(m)}(t, \kappa)$ represents the generalized time-dependent force acting on the mode (m, κ), and $Q^{(m)}(t, \kappa)$ is the time-dependent response of the amplitude of the mode. The total plate response in (5.1) is represented as a superposition of the response of each normal mode. The choice of receiver location and the choice of loading here find their expression through the "receiver function" $\mathbf{u}_m$ and through the "source function" Q.

It is particularly revealing to note that the modal factor depends on the source character and location via the integral of $F^{(m)}$ in equation (5.3). Thus if the force $\mathbf{F}(r, z, t)$ is orthogonal to a mode $\mathbf{u}^{(m)}(r, z, \kappa)$, or more simply, if the simple point force acts at a node of the mode in question, that mode is not excited by the force; analogously, a string plucked at its midpoint is not excited into any even harmonics.

5.2 Waves Generated by a Stationary Concentrated Force. For the remaining part of this paper, we consider a vertical, stationary, concentrated force at $(0, 0, z_0)$, the time dependence being a Heaviside step function. We thus express $\mathbf{F}(\mathbf{x}, t)$ as

$$\mathbf{F}(\mathbf{x}, t) = Q_0\frac{\delta(r)}{2\pi r}\delta(z-z_0)H(t)\,\hat{\mathbf{z}} \tag{5.5}$$

where Q_0 is the magnitude of the applied force. Substituting it into equation (5.3), we find

$$F^{(m)}(t, \kappa) = H(t) f_m(z_0, \kappa) \tag{5.6}$$

where, for antisymmetric modes,

$$f_m(z_0, \kappa) = Q_0\kappa^2[2\alpha_m\beta_m \sin\beta_m h \cos\alpha_m z_0 + (\kappa^2-\beta_m^2)\sin\alpha_m h \cos\beta_m z_0] \quad m = 1, 3, 5, \ldots \tag{5.7a}$$

and for symmetric modes,

$$f_m(z_0, \kappa) = Q_0\kappa^2[2\alpha_m\beta_m \cos\beta_m h \sin\alpha_m z_0 + (\kappa^2-\beta_m^2)\cos\alpha_m h \sin\beta_m z_0] \quad m = 2, 4, 6, \ldots, \tag{5.7b}$$

The time function in equation (5.2) can then easily be integrated,

$$Q^{(m)}(t, \kappa) = f_m(z_0, \kappa)\frac{1}{\omega_m^2(\kappa)}[1-\cos\omega(\kappa)t]H(t) \tag{5.8}$$

To summarize, the final result for transient waves in an infinite plate with thickness $2h$, which is excited by a concentrated vertical force being placed suddenly at $t = 0$, $r = 0$, and $z = z_0$ is

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\mathbf{u}(r, z, t) =$$

$$H(t) \sum_{m=1}^{\infty} \int_0^{\infty} f_m(z_0, \kappa) \frac{\mathbf{u}_m(r, z; \kappa)}{\pi M_m(\kappa)} \left[\frac{1 - \cos \omega_m(\kappa) t}{\omega_m^2(\kappa)} \right] d\kappa. \quad (5.9)$$

In the foregoing, the $f_m(\kappa)$ are given in equation (5.7), and the modes for the mth branch are

$$\mathbf{u}_m(r, z; \kappa) = U_m(r, z; \kappa)\hat{\mathbf{r}} + W_m(r, z; \kappa)\hat{\mathbf{z}} \quad (5.10)$$

where the U_m and W_m functions are given in equations (3.9) and (3.10). The $M_m(\kappa)$ equal $M_\infty^{(m)}$ in equation (5.1) and are given by equation (4.20). The ω_m are related to κ by equation (5.4). These equations form the basis for the stationary phase analysis and for the numerical evaluations.

6 Transient Response by Integral Transforms

The standard procedure for the analysis of the axisymmetric response of a plate to dynamic loadings entails the performance of two integral transforms on the equations of motion, one on the time variable, the other on a spatial variable. This procedure is thoroughly discussed by Miklowitz (Chapter 7, reference [4]) for a line load and plane strain. We will not present the analogous procedure for the axisymmetric case but will describe it and then compare it with the method of eigenfunction expansion.

For the double transform technique equation (2.1) is re-expressed as decoupled scalar equations for the displacement potentials ϕ and ψ involving the force potentials f and g, where

$$\mathbf{u} = \nabla\phi + \nabla \times \nabla \times (\psi\hat{\mathbf{z}}) \quad (6.1)$$

$$\mathbf{F} = \nabla f + \nabla \times \nabla \times (g\hat{\mathbf{z}}). \quad (6.2)$$

Similarly equations (2.2) are rewritten as coupled boundary conditions on ϕ and ψ. A Fourier (or Laplace) transform in time and a Hankel transform in radial position r are performed on these rewritten equations. The resulting coupled inhomogeneous ordinary differential equations for the double transforms of ϕ and ψ may be solved and the inverse double transforms performed to obtain the actual displacement potentials. $\mathbf{u}$ is then constructed from (6.1).

If the inverse Fourier transform is performed before the inverse Hankel transform and evaluated by means of contour integration and the theory of residues, the resulting expression for $\mathbf{u}$ resembles (5.9), and must necessarily be identical to equation (5.9). Alternatively, one may do the inverse Hankel transform first, also by residues, and directly obtain the spectrum of the response. The transient solution would then be obtained by numerical inverse Fourier transformation.

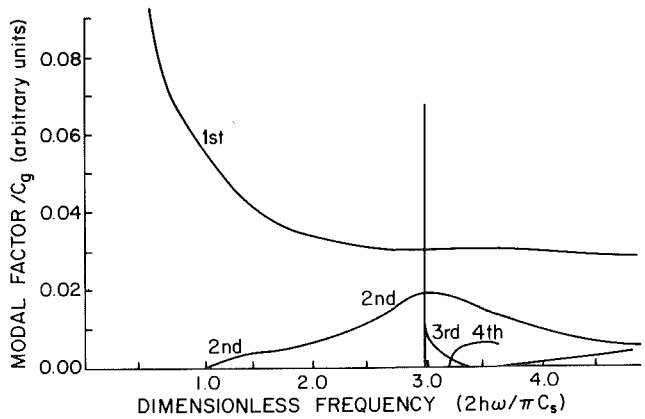


Fig. 5(a) Antisymmetric modes

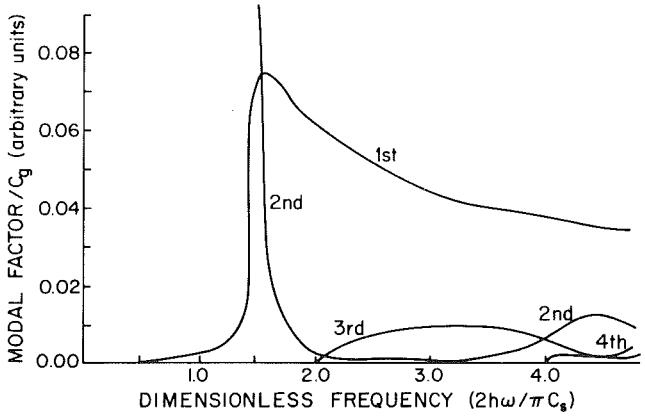


Fig. 5(b) Symmetric modes

Fig. 5 Modal factor/group velocity

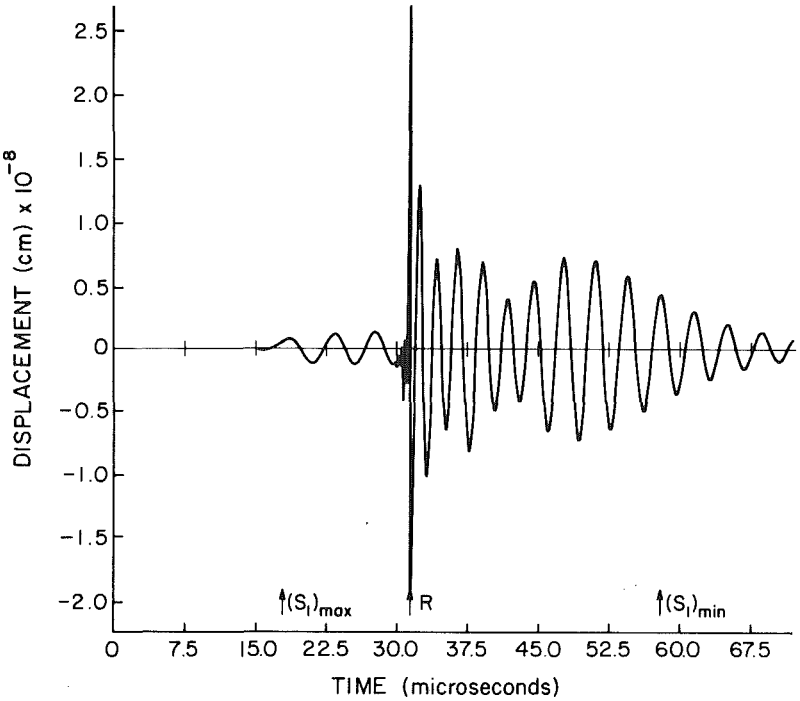


Fig. 6(a) First symmetric branch (S_1)

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

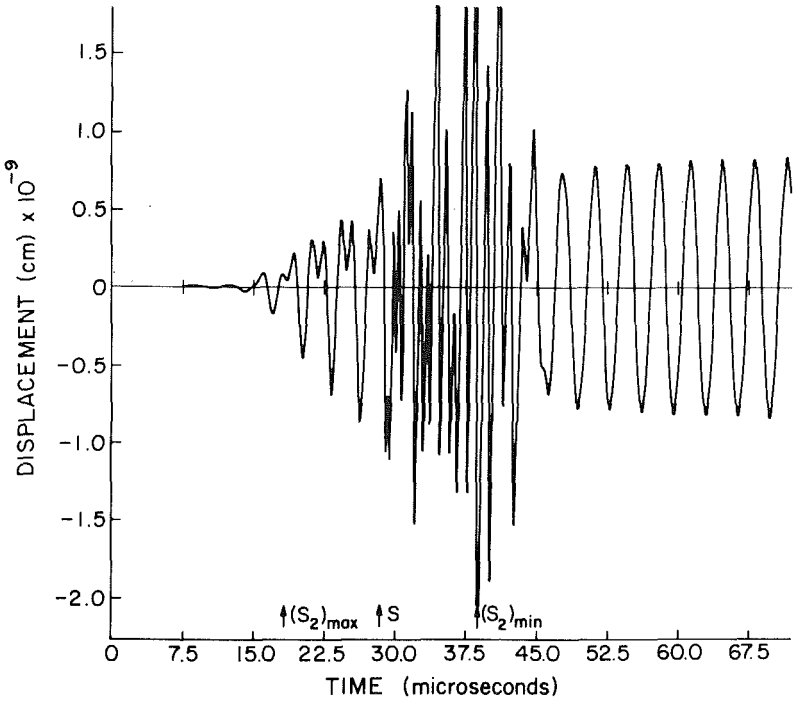


Fig. 6(*b*) Second symmetric branch (S_2)

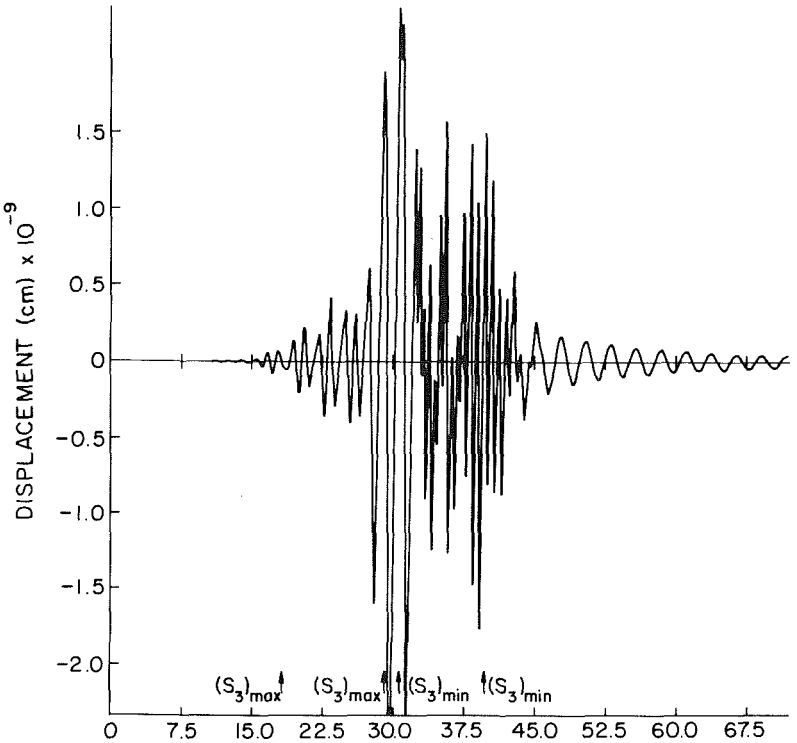


Fig. 6(c) Third symmetric branch (S_3)

Fig. 6 Normal displacement due to individual symmetric branches ($r = 10$ cm)

One can with reason claim that the method of integral transforms is more versatile in that with it one has two options of evaluating the inverse transforms. One option leads directly to the solution by the method of normal modes, and the other yields the same transient solution as well as the spectrum of the transient response. The practical value of this aspect of the double transform method is, however, limited, because evaluation of the spectrum requires the location of all complex roots of the Rayleigh-Lamb function as well as the real roots. Furthermore, for buried forces the method of integral transforms requires the decomposition of the force into force potentials f and g which is tedious, if not very difficult especially in the case of nonaxisymmetric loads. The method then requires the solutions of the ordinary differential equations in z. These solutions are very complex in the general case of buried distributed forces. Miklowitz's applications of the integral transform method have been confined to those loads representable by inhomogeneous boundary conditions

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

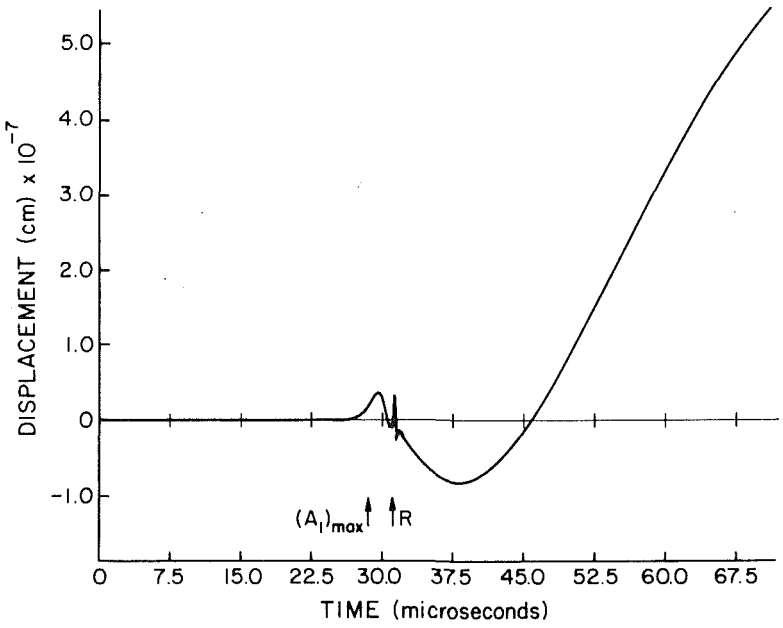


Fig. 7(*a*) First antisymmetric branch (A_1)

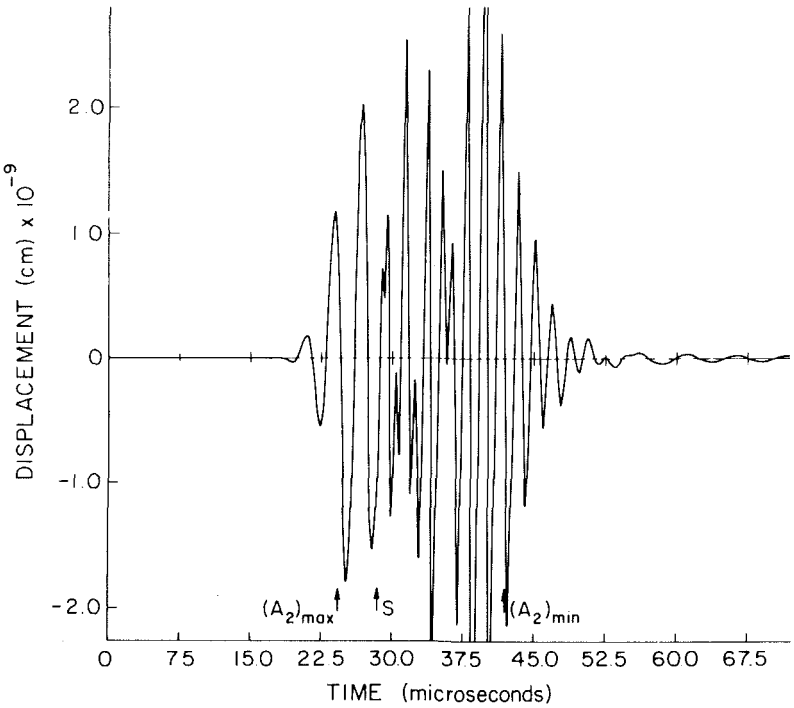


Fig. 7(*b*) Second antisymmetric branch (A_2)

Fig. 7 Normal displacement due to individual antisymmetric branches (r = 10 cm)

such as vertical surface forces where the procedure is much simplified.

It is our contention that the integral transform method is analytically the more cumbersome and, for many applications, the less versatile. It can further be criticized on the grounds that the use of the abstract quantities, the transforms of the displacement potentials, obscures the simple physics which is revealed in the method of explicit expansion in normal modes.

7 Analysis by the Method of Stationary Phase

Analysis of integrals such as equation (5.9) has in the past been confined essentially to far field asymptotic approximations based on the method of stationary phase. One recognizes that at large r, ($r >> h$ or $t >> h/c_p$) the variation of the integrand with κ is dominated by the rapid oscillations of the Bessel functions $J_0(\kappa r)$ and $J_0'(\kappa r)$ in $\mathbf{u}_m$ and the function cos $\omega_m(\kappa)t$. But if $\kappa = \kappa_0$ is such that $|\partial\omega/\partial\kappa|_{\kappa=\kappa_0} = r/t$ then the oscillations of the cosine function and the Bessel functions are in phase, whereupon in the vicinity of κ_0 the integrand is, for a short range in κ, not rapidly oscillating. The displacement time trace at r is thus dominated by a sum of terms oscillating in time; each of them represents a simple harmonic wave train at a frequency such that $|\partial\omega/\partial\kappa| = r/t$. This is the basis for the well-known "group velocity analysis" [22] or "period-time of occurrence relations" [4]. The relative strengths of these many contributions depends of course on the magnitudes of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$f_m(\kappa)/M_m(\kappa)$ when evaluated at $\kappa = \kappa_0$. This analysis shows further that each such contribution is inversely proportional to $(t|\partial^2\omega/\partial\kappa^2|)^{1/2}$ and thus the points at which the group velocity is nearly stationary are expected to be the most important. The actual point where $\partial^2\omega/\partial\kappa^2 = 0$ in fact contributes an amount proportional to $(t\partial^3\omega/\partial\kappa^3)^{-1/3}$ and thus dominates at sufficiently large t [23]. The contributions from such points are called the "Airy phases."

As an illustration, we consider the normal displacement u_z at r and $z = +h$, the force Q_0 of equation (5.5) being applied at the surface of the plate ($z_0 = +h$),

$$u_z(r, h, t) = H(t) \sum_{m=1}^{\infty} \int_0^{\infty} f_m(h, \kappa) \frac{W_m(r, h; \kappa)}{\pi M_m(\kappa)} \frac{1 - \cos \omega_m(\kappa) t}{\omega_{m(\kappa)}^2} d\kappa \qquad (7.1)$$

We separate out the r-dependence of W_m of equation (5.10), or equations (3.9) and (3.10), and rewrite the previous result as

$$u_z(r, h, t) = H(t) \sum_{m=1}^{\infty} u_{zm}(r, h, t)/\pi \qquad (7.2)$$

where

$$u_{zm}(r, h, t) = \int_0^{\infty} Z_m(h, \kappa) J_0(\kappa r)(1 - \cos \omega_m t)\, d\kappa \qquad (7.3)$$

and

$$Z_m(h, \kappa) = f_m(h, \kappa) W_m(r, h; \kappa)/[M_m(\kappa)\omega_m^2(\kappa) J_0(\kappa r)]$$

The integration may be separated into two parts. The time-independent part yields the "statical deflection" for each branch corresponding to the long-time contribution of the step loading. The time-dependent part can be evaluated by the method of stationary phase, by noting that for large t and r,

$$J_0(\kappa r)\cos \omega_m t \sim (2\pi\kappa r)^{-1/2} Re[e^{i(\kappa r - \omega_m t - \pi/4)} + e^{i(\kappa r + \omega_m t - \pi/4)}] \qquad (7.4)$$

Along each branch of the group velocity curves (Fig. 3), there might be more than one maximum and minimum point. Corresponding to each stationary point, there are a set of values for the wave number, frequency, and group velocity, which are designated as κ_0, ω_0, c_{g0} respectively, where

$$c_{g0} = (\partial\omega/\partial\kappa)_{\kappa=\kappa_0} \qquad (7.5)$$

Since the arrival time of this wave-train group at r is $t_0 = r/c_{g0}$, the group with the maximum speed will be well separated from the group with the minimum speed on the time trace recorded at r. For time t around t_0, the dominant value of the time-dependent part of the integral, equation (7.3), is, [22] the contribution from the modes near the stationary point in the group velocity curve, the "Airy phase."

$$u_{zm} \sim (2\pi\kappa_0 r)^{-1/2} Z_m(h, \kappa_0)(t\omega_0'')^{-1/3} \mathrm{Ai}(\tau)\cos(\omega_0 t - \kappa_0 r + \pi/4) \qquad (7.6)$$

where

$$\tau \equiv c_{g0}(t - t_0)(t\omega_0''/2)^{-1/3} \qquad (7.7)$$

In the foregoing, $\omega_0'' = (d^3\omega/d\kappa^3)_{\kappa=\kappa_0}$ and Ai is the Airy function. The preceding time trace may be described as a rapidly varying sinusoidal function of time with a slowly varying envelope of an Airy function. The strength of this contribution to the total displacement depends of course on the magnitude of the function $Z_m(\kappa)$ at the point κ_0.

Typical plots of the foregoing expression are given in Figs. 4 where the factor $Z/(2\pi\kappa_0 r)^{1/2}$ has been replaced by unity and where the plate parameters have been chosen, $\nu = 0.21$, $2h = 0.960$ cm, and $c_s = 0.349$ cm/μsec. The receiver is located a distance of 40 thicknesses from the source. Figure 4(a) shows the case $\omega_0'' < 0$, i.e., the case of a maximum group velocity for the second antisymmetric branch at $\omega_0 = 1.805\ \omega_s$ ($\omega_s = \pi c_s/2h$), $\kappa_0 = 0.94\ (\pi/2h)$, $c_{g0} = 1.17\ c_s$ and $\omega_0'' = -1.34\ (2h/\pi)^2\ c_s$. Figure 4($b$) shows the case of the minimum group velocity for the first symmetric branch at $\omega_0 = 1.55\ \omega_s$, $\kappa_0 = 1.24\ (\pi/2h)$, $c_{g0} = 0.49\ c_s$, and $\omega_0'' = 3.37\ (2h/\pi)^2\ c_s$. The former case may be called a "mode arrival," the latter a "mode disappearance." As shown in the foregoing one might expect curves of these forms to dominate the time traces for sufficiently large source-receiver separations. Wave packets with these shapes will be found in the general results shown in the next section.

Stationary phase analysis of integrals such as the preceding has been employed at least since the time of Kelvin (1887) and indeed is the only analysis possible in the far field short of extensive numerical work. Practical difficulties exist with it however, chiefly describable by the following unanswered questions: How large must r and t be in order that the Airy phases are the major contributions? What are the appropriate expressions, different from equation (7.6), at those maxima and minima that exist at the end points, $\kappa = 0$ or $\kappa = \infty$? Furthermore there exist points where $\omega'' \approx \omega''' \approx 0$ where expression (7.6) is again invalid. There also exist points where $\omega'' = 0$ and $Z_m \approx 0$, where again the preceding expression is invalid.

For convenience, we call $Z_m(h, \kappa)$ of equation (7.3) the "modal factor." It depends on the loading function, $\omega_m^{-2} f_m(h, \kappa)$, as well as the normal mode, $U_m(r, h; \kappa)/J_0(\kappa r)$. It may become zero or infinite at certain values of κ, a point not considered in the usual problems of stationary phase analyses. In Fig. 5 we show $Z_m(h)$ (divided by $|c_g| = |d\omega/d\kappa|$) as a function of ω/ω_s for the first four branches of antisymmetric modes (Fig. 5(a)), and first four symmetric branches (Fig. 5(b)). These curves are dependent on the vertical location of the source ($z_0 = h$) and receiver ($z = h$), the time function of the concentrated force (a step time function), but independent of the radial distance r and time t. It is interesting to note that values of the modal factor are zero, or nearly zero, at the maxima of group velocities of many branches (Fig. 3). These modes are thus much less significant then otherwise have been expected.

The conceptual complexities required of a thorough pursuit of this method are many. Thus in the light of its approximate nature, we suggest that its value is chiefly for qualitative analysis, and in the initial interpretation of measured plate responses. This qualitative theory may be compared to the numerical results of the exact integral as given in the next section.

8 Surface Response of a Plate Due to a Concentrated Force

The integrals in equations (7.2) and (7.3) for the vertical displacement component, u_z, have been evaluated by a numerical integration. The algorithm employed for the numerical integration entailed a conversion of independent variable from κ to ω and then an integration by Fast Fourier Transform techniques (FFT). Because the conversion of independent variable introduces an inverse factor of group velocity to the integrand and because FFT techniques do not allow for varying step sizes in the integration, the low group velocity ranges are not accurately computed by this technique. Integration for these ranges was done directly in the wave number domain with a trapazoidal rule. This algorithm was thought to be the most efficient and accurate for a not too large range in source-receiver distances. At very large distances where the integrand is in general a rapidly varying function of κ (or ω), the integral would probably be most efficiently and accurately evaluated by a method appropriate to the numerical integration of rapidly and sinusoidally varying functions [24]. The computing costs varied with the source-receiver distance but averaged about 1 dollar per integration.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

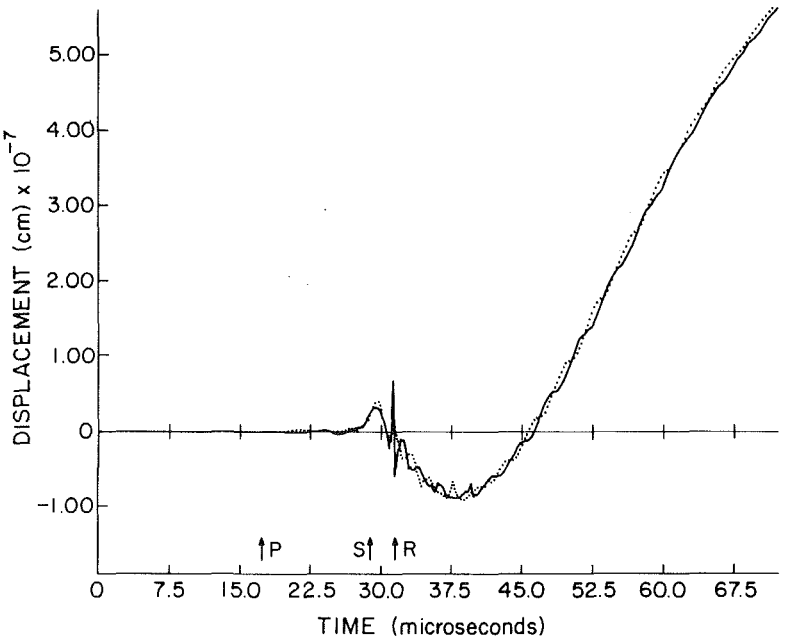


Fig. 8(*a*) The long time response at upper ($z = h$) and lower surface ($z = -h$)

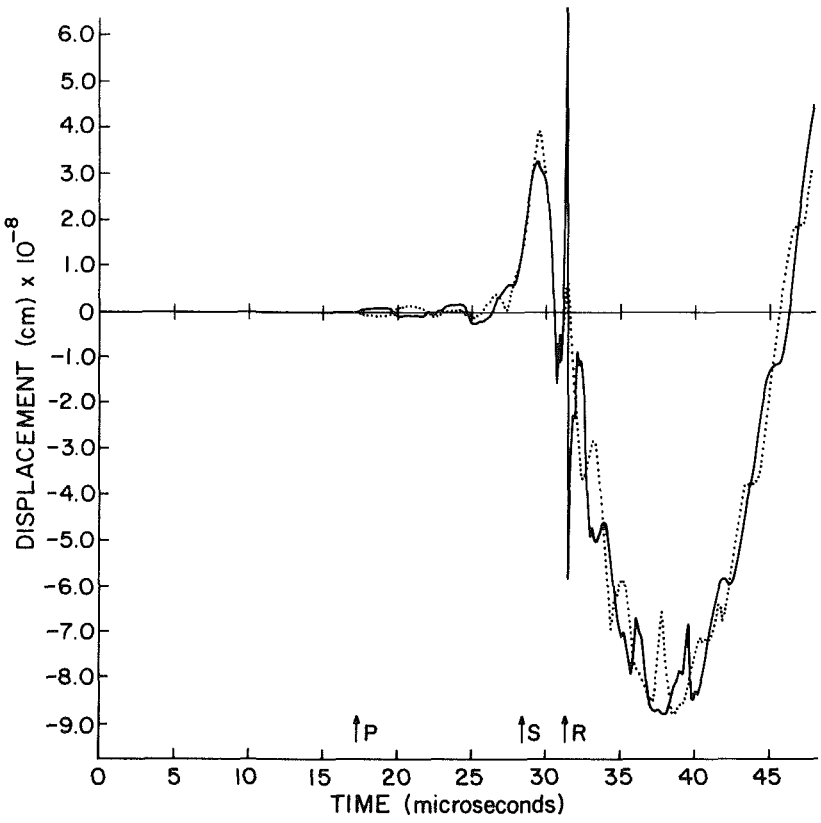


Fig. 8(*b*) Early time response in expanded scale

Fig. 8 Vertical displacements at the surfaces of the plate, solid line for $z = h$, dotted line for $z = -h$ ($r = 10$ cm)

The ensuing plots are the calculated vertical displacement responses of the surface of a glass plate due to an excitation of a step vertical downward force of 352 g; that is $Q_0 = -3.46$ N in equation (6.5),

$$\mathbf{F} = -346{,}000 \text{ (dynes)}\hat{\mathbf{z}}\,\delta(r)\delta(z-h)H(t)/2\pi r.$$

The glass has Poisson ratio 0.21 which is calculated from a shear wave speed $c_s = 0.349$ cm/μsec and a pressure wave speed $c_p = 0.576$ cm/μsec. Material density is 2.3 g/cm^3, and thickness of the plate is 0.96 cm. The values of these parameters were chosen to match the parameters of experimental work being done by Sachse and Pao [25].

8.1 Contributions From Individual Branches. In Figs. 6 and 7 we show u_{zm} (r, h, t) of equation (7.3) at a distance of 10 cm on the top surface of the plate ($z = h$) due to some of the lower branches. These plots can be qualitatively understood by the stationary phase analysis of the preceding section.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

One clearly distinguishes, for instance, three major features in Fig. 6(*a*), the contribution of the first symmetric branch ($m=2$). One feature is the early arrival of a relatively low-amplitude and low-frequency component. The next feature is the sharp spike of very high frequency arriving at about 31 μsec. Finally there is, in the neighborhood of 60 μsec, an apparently exponentially decaying harmonic wave of a single frequency.

The plot of group velocity versus frequency for the first symmetric branch (Fig. 3(*a*)) shows three stationary points. The first is at a large group velocity $c_g \cong c_p$ and at low frequency. Associated with the point is a very low value for the modal factor (Fig. 5(*b*)). A group moving at this speed would arrive at $t = 18.0$ μsec in Fig. 6(*a*) (marked by $(S_1)_{max}$). There is also a local maximum at infinite frequency, with $c_g = c_R = 0.319$ cm/μsec associated with a nonzero modal factor. This occurs at R in Fig. 6(*a*) as the sharp Rayleigh peak ($t = 31.4$ μsec). Technically this peak ought to be of infinite height. A high-frequency cutoff in the numerical work where the modal factor Z has been set smoothly to zero above 2 MHz, has eliminated the singularity. Finally there is a minimum in the group velocity at a moderate frequency. Thus one expects and finds a late contribution Airy phase like that in Figure 4(*b*), tailing off at the point $(S_1)_{min}$ in Fig. 6(*a*) at $t = 58.0$ μsec. Each of these features is qualitatively predictable by the methods of Section 7.

The plot of displacement of the second symmetric branch S_2 in Fig. 6(*b*) also shows three major features. One is a nearly

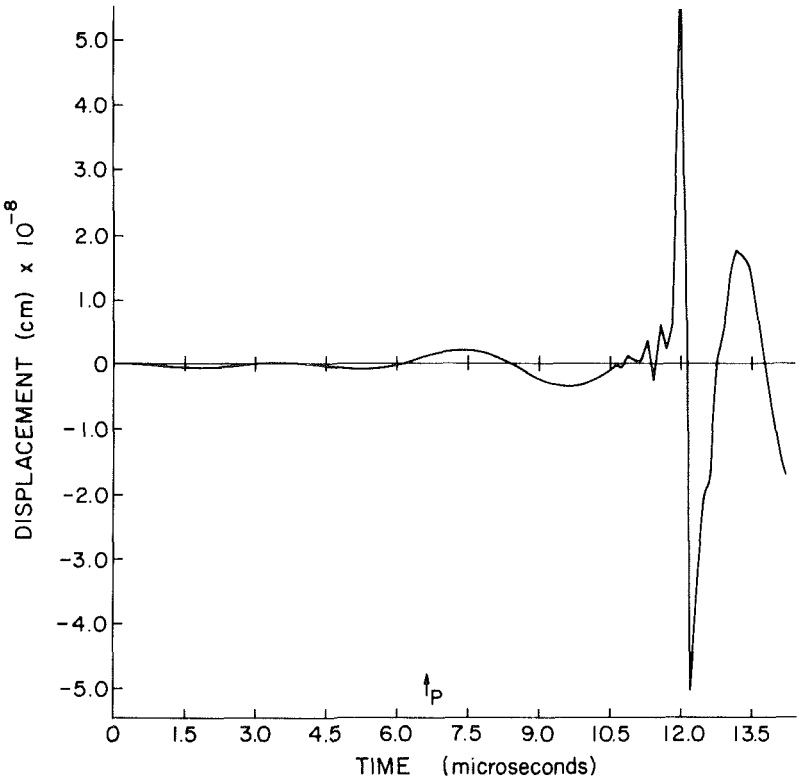


Fig. 9(*a*) Noncasual behavior due to the first symmetric branch

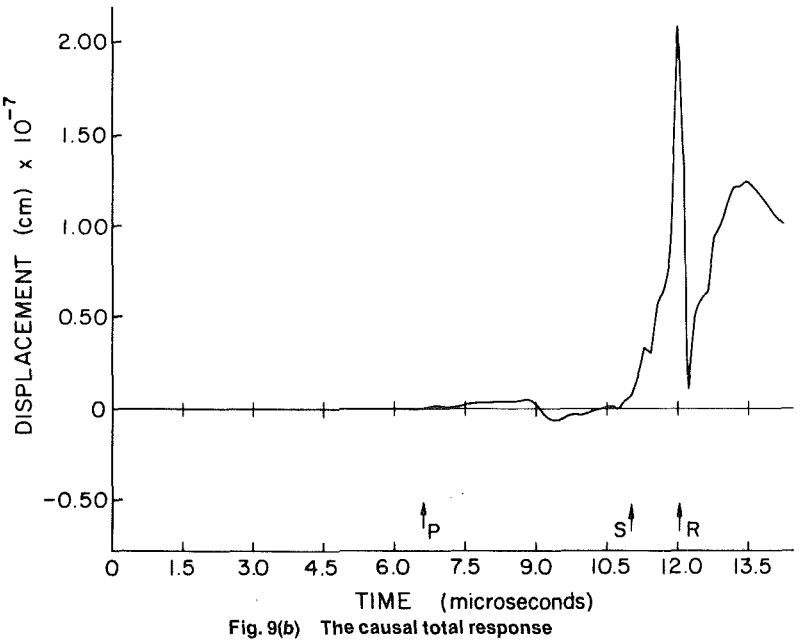


Fig. 9(*b*) The causal total response

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

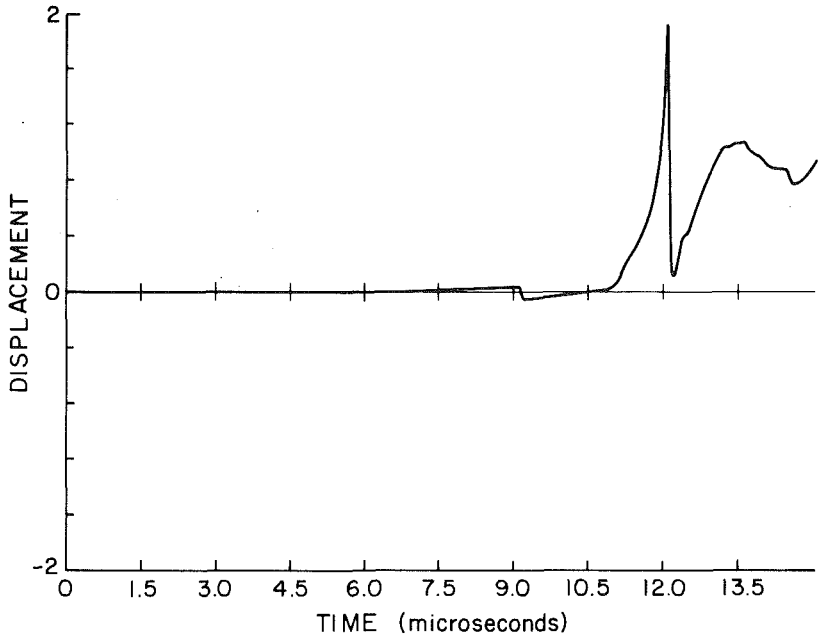


Fig. 9(c) Exact result of generalized ray theory

Fig. 9 Normal displacement at the upper surface of the plate ($r = 8h$)

constant frequency, very late large-amplitude component. Another is a high-frequency component confined to the times 30 μsec–43 μsec. The third is an early arrival of relatively low amplitude. These features also can be understood by considering the group velocity curve in Fig. 3(b) and the modal factor in Fig. 5(b). The long time scale ringing for instance is due to the large amplitude, low group velocity region near the cutoff frequency at $\kappa = 0$, and $\omega = 1.65\ \omega_s$.

The results shown in Fig. 6(c) for the third symmetrical branch, S_3, are similar to those of S_2.

The lowest antisymmetric branch (Fig. 7(a)) is of particular interest. It also has three main features. The first is a moderately low-frequency arrival approximately at the shear wave speed. There is also a sharp peak at the Rayleigh speed much like the same peak in Fig. 6(a). Finally there is an extremely low-frequency high-amplitude disturbance at late times. The sharp peak marked by R is a Rayleigh wave corresponding to the high-frequency asymptotic minimum in the group velocity. The early arrival is due to the sole maximum in group velocity at $(c_g)_{\max} = 0.996\ c_s$. The very large amplitude low-velocity disturbance corresponds to modes for which the group velocity is not stationary but for which the modal factor is immense. This is the low-frequency regime of flexural motion where the Germain-Lagrange bending theory would be valid [26].

Wave motion contributed by the second antisymmetric branch (A_2, $m = 3$) is shown in Fig. 7(b). Its features can be interpreted, as in the case for S_2, by the group velocity curve in Fig. 3(a) and modal factor curve in Fig. 5(a). The arrival times of various Airy phases $(A_2)_{\max}$, $(A_2)_{\min}$, and the arrival time of a shear ray are marked on the figure.

8.2 Normal Displacement at 10 Centimeters. Finally, in Figs. 8(a) and 8(b) we show the vertical displacement at 10 cm on the top at $z = h$ (solid lines), and on the bottom at $z = -h$ (dotted lines). The sum in equation (5.9) has been truncated at the first 10 branches (m from 1–10), thus fairly accurately including effects up to a frequency of the cutoff for the 11th branch ($\omega = 6\omega_s$, $f = 1.09$ MHz). As the modal factors appear to be dropping steadily with increasing order of branches, the indicated "total" displacement can be expected to be fairly good even above 1.09 MHz.

The upward displacement on top is of course the sum of contributions from each of the 10 branches at $z = h$ as discussed in the preceding subsection. The upward displacement at $z = -h$ is the sum of all the antisymmetric contributions at $z = h$ *minus* the sum of all the symmetric contributions at $z = h$. Thus, amongst other effects, the sharp Rayleigh wave peak due to the symmetric and antisymmetric lowest branches appears on the top surface, but does not appear on the bottom surface. To show the details of the Rayleigh arrival, the early time response at both surfaces are enlarged in Fig. 8(b).

8.3 Comparison With Response Based on Generalized Ray Theory. A particularly interesting phenomenon is exhibited in Fig. 9(a) for normal displacement at $r = 8h = 3.84$ cm due to the first symmetric branch. The response due to this branch alone is noncasual in the sense that there is disturbance before the arrival time of the P ray at 6.67 μsec. Similar noncausal behavior also exists for higher symmetric branches. One is thus suspicious of numerical error. The noncausal result for individual branches is, however, real, because the total response from all branches can still be causal (Fig. 9(b)).

Within the resolution of the numerical data, the noncausal parts of the second, third, and to a lesser extent, the higher symmetric branches have destructively interferred with those of the first branch, resulting in a correct result with zero response before the arrival of the P-wave. The sum of the contributions from the first 10 branches, a causal response, is shown in Fig. 9(b). The noncausal behavior for individual branches is also present at $r = 10$ cm but is much less pronounced there. The noncausual results, though initially surprising, cause no conceptual difficulties since a "branch" as discussed in Section 2 is not a concept fundamental to the plate response. It is impossible in principle to either excite or build a receiver receptive to only modes on a given branch. Similar results are reported in a treatment of elastic wave scattering (section 8.3, reference [4]).

In Fig. 9(c) we also show the exact result based on the generalized ray theory [11]. Its agreement with the result of normal mode theory appears good except for frequency components of the order of a megahertz or higher, as expected.

8.4 Response at 80h. The total surface response at $r = 80\ h$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

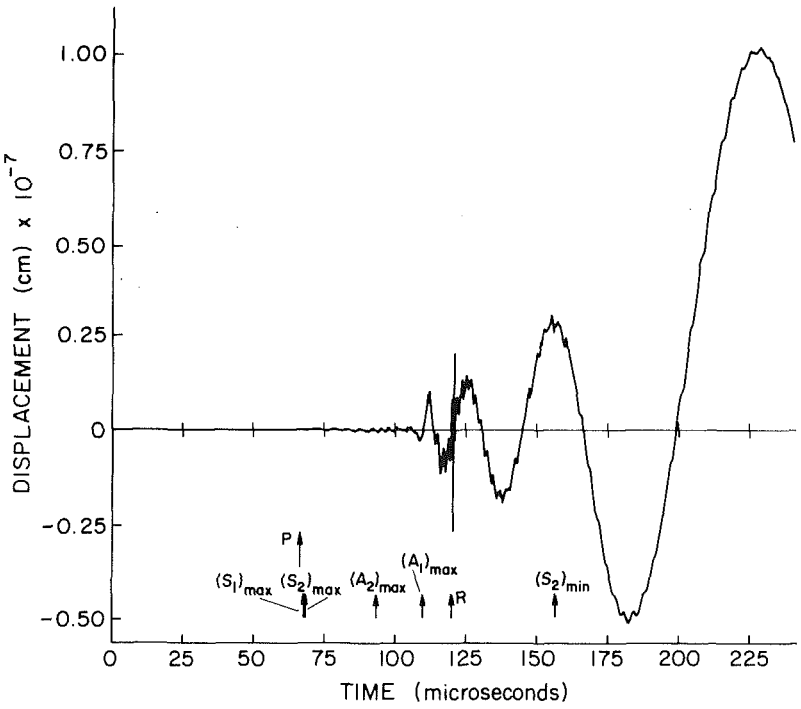


Fig. 10(*a*) Surface response at $r = 80\,h$

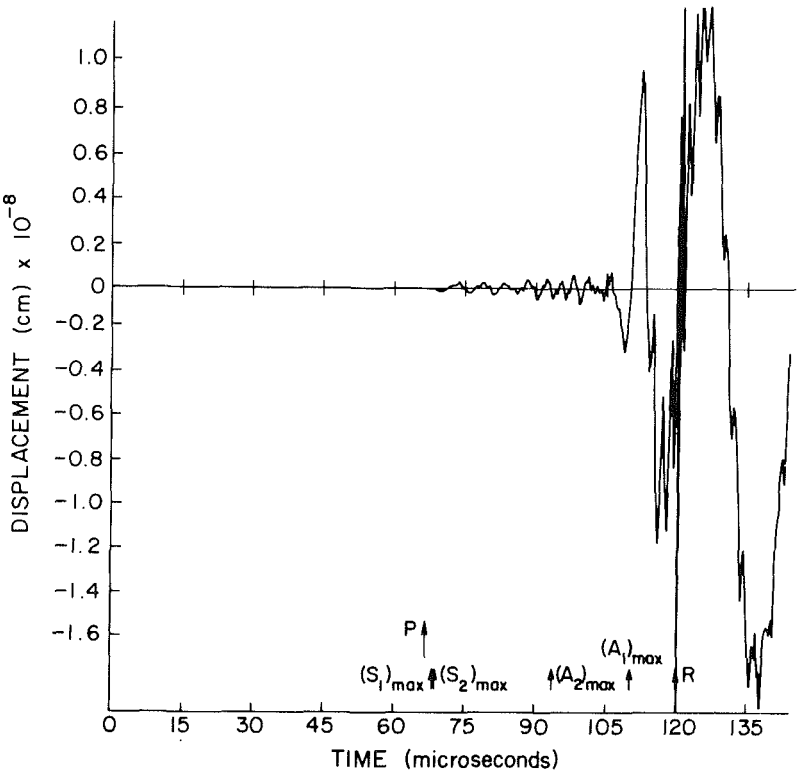


Fig. 10(*b*) Normal displacement at $z = +h$ in expanded scale

Fig. 10 Surface response at $r = 80\,h$

is calculated from a sum of 10 branches and is shown in Figs. 10(*a*) and 10(*b*). The distance is in the "intermediate range" ($20\,h < r < 200\,h$), and some of the features of the group velocity analysis should be identifiable. The arrival times of Airy phases traveling at some of the maximal and minimal group velocities are marked in the figure.

The early part of the displacement at $z = +h$ of Fig. 10(*a*) is shown enlarged in Fig. 10(*b*). In both figures, the modes of $(S_1)_{max}$ and $(S_2)_{max}$ are, however, not clearly present because the corresponding modal factors are nearly zero. The late arriving groups such as $(S_2)_{min}$, and the groups near the branch cutoff frequencies, such as the part responsible for the long time-scale ringing in Fig. 6(*b*), are somewhat masked in Fig. 10(*a*) by the large-amplitude, low-velocity, low-frequency, simple flexural response of the first antisymmetric branch. The presence of the higher frequency late-arriving groups would be more apparent in the output of a transducer with poor low-frequency response. There the group arrivals

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

would be somewhat more obvious, as would the long time-scale ringing, due predominently to the cutoff frequency domain of the second symmetric branch.

Comparison of these results with experimental results is problematic. Presently available experimental responses at large distances and long times are complicated by the presence of an unknown transducer transfer function [27]. Thus the voltage output of a pizoelectric transducer should not be compared directly with the displacement curves given here. Nevertheless a comparison has been made and the differences reasonably explained by some unknown linear transducer transfer function. Further work along these lines is ongoing.

9 Conclusion

In this paper, we have applied the method of expansion in normal modes to the solution of transient excitation of a thick plate, first to a plate with finite radius, then to an unbounded plate. The method of normal modes is chosen instead of the usual method of integral transform because it treats a buried force or moving force as easily as a surface force.

The final results for plate displacement are expressed as a series of integrals, each being integrated over the radial wave numbers of one branch of the Rayleigh-Lamb spectrum. These integrals are evaluated numerically up through the tenth branch.

Numerical results for the surface response of a plate at 10.42, 4, and 40 plate thickness from the source are shown in Figs. 8, 9(*b*), and 10, respectively, the source being a vertical force suddenly applied at the surface of the plate. Some features of these responses are interpreted in terms of a stationary phase analysis and the group velocities of Rayleigh-Lamb waves in the plate. Within the very large present experimental uncertainties, the result at 40 plate thickness is in agreement with the experimental measurement. The result at 4 plate thicknesses agrees very well with the exact results derived from the theory of generalized ray. We conclude that the method of normal modes can be applied effectively to analyze transient progressive waves in a plate at intermediate and far range.

Acknowledgment

The authors acknowledge the support provided by the National Science Foundaton through a grant to the College of Engineering (ENG 7910311), and one to the Materials Science Center of Cornell University.

References

1 Lord Rayleigh, "On the Free Vibrations of an Infinite Plate of Homogeneous Isotropic Elastic Matter," *Proceedings London Mathematical Society*, Vol. 20, 1888–1889, p. 225.

2 Lamb, H., "On Waves in an Elastic Plate," *Proceedings Royal Society London*, Series A, Vol. 93, 1916–1917, p. 114.

3 Mindlin, R. D., "Waves and Vibrations in Isotropic, Elastic Plates," in: *Structural Mechanics*, Goodier, J. N., and Hoff, N. J., Eds., Pergamon Press, New York, 1960.

4 Miklowitz, J., *The Theory of Elastic Waves and Wave Guides*, North-Holland, Amsterdam, 1978.

5 Lloyd, J. R., and Miklowitz, J., "On the Use of Double Integral Transforms in the Study of Dispersive Elastic Wave Propagation," *Proceedings of the Fourth U.S. National Congress of Applied Mechanics*, Vol. 1, The American Society of Mechanical Engineers, New York, 1962, pp. 255–267.

6 Miklowitz, J., "Transient Compressional Waves in an Infinite Elastic Plate or Elastic Layer Overlying a Rigid Half-Space," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, 1962, pp. 53–60.

7 Scott, R. A., and Miklowitz, J., "Transient Nonaxisymmetric Wave Propagation in an Infinite Isotropic Elastic Plate," *International Journal of Solids and Structures*, Vol. 5, 1969, pp. 65–79.

8 Mencher, A. G., "Epicentral Displacement Caused by Elastic Waves in an Infinite Slab," *Journal of Applied Physics*, Vol. 24, 1953, pp. 1240–1246.

9 Knopoff, L., "Surface Motions of a Thick Plate," *Journal of Applied Physics*, Vol. 29, 1958, pp. 661–670.

10 Davids, N., "Transient Analysis of Stress-Wave Penetration in Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 26, 1959, p. 651.

11 Pao, Y. H., Gajewski, R. R., and Ceranoglu, A. N., "Acoustic Emission and Transient Waves in an Elastic Plate," *Journal of the Acoustical Society of America*, Vol. 65, 1979, pp. 96–105.

12 Ceranoglu, A. N., and Pao, Y. H., "Propagation of Elastic Pulses and Acoustic Emission in a Plate, Part I. Theory, Part II. Epicentral Response, Part III. General Responses," ASME JOURNAL OF APPLIED MECHANICS, 1981, pp. 125–147.

13 Love, A. E. H., *Mathematical Theory of Elasticity*, 4th ed., Dover, New York, 1944.

14 Eringen, A. C., and Suhubi, E. S., *Elastodynamics, Vol. 2, Linear Theory*, Academic Press, New York, 1975.

15 Reismann, H., "On the Forced Motion of Elastic Solids," *Appl. Sci. Res.*, Vol. 18, 1967, p. 156.

16 Goodman, L. E., "Circular Crested Vibrations of an Elastic Solid Bounded by Two Parallel Planes," *Proceedings of First U.S. National Congress of Applied Mechanics*, The American Society of Mechanical Engineers, New York, 1951, p. 65.

17 Meeker, T. R., and Meitzler, A. H., "Guided Wave Propagation in Elongated Cylinders and Plates," in: *Physical Acoustics*, Vol. 1A, Mason, W. P., Ed., Academic Press, New York, 1964, pp. 111–167.

18 Pao, Y. H., and Kaul, R. K., "Waves and Vibrations in Isotropic and Anisotropic Plates," in: *R.D. Mindlin and Applied Mechanics*, Herrmann, G., Ed., Pergamon Press, New York, 1974.

19 Hutchinson, J. R., "Vibrations of Solid Cylinders," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 901–907.

20 National Bureau of Standards, *Handbook of Mathematical Functions*, Abramowitz, M., and Stegun, I. A., Eds., U.S. Government Printing Office, Washington, D.C., 1964.

21 Pursey, H., "The Launching and Propagation of Elastic Waves in Plates," *The Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 10, 1957, p. 45.

22 Copson, E. T., *Asymptotic Expansions*, Cambridge University Press, Chapter 4, 1965.

23 Chester, C., Friedman, B., and Ursell, F., "An Extension of the Method of Steepest Descent," *Proceedings Cambridge Philosophical Society*, Vol. 53, 1957, p. 599.

24 Longman, I. M., "Note on a Method for Computing Infinite Integrals of Oscillating Functions," *Proceedings Cambridge Philosophical Society*, Vol. 52, 1956, p. 764.

25 Sachse, W., and Pao, Y. H., "Propagation of Wide Band Pulses in an Elastic Plate," in preparation.

26 Medick, M., "On Classical Plate Theory and Wave Propagation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, 1961, p. 223.

27 Sachse, W., and Hsu, C. S., "Ultrasonic Transducers for Materials Testing and Their Characterization," in: *Physical Acoustics*, Vol. 14, Mason, W. P., and Thurston, R. N., Eds., Academic Press, New York, 1979, pp. 277–406.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Y. J. Chao
Department of Engineering Mechanics and Materials, Southern Illinois University at Carbondale, Carbondale, Ill. 62901

M. A. Sutton
Department of Mechanical Engineering, University of South Carolina, Columbia, S. C. 29208

C. E. Taylor
Department of Engineering Science, University of Florida, Gainesville, Fla. 32601

Interferometric Determination of Curvatures of Flexed Plate

Two interferometric techniques are proposed that can readily be applied to the measurement of stresses in thin plates. A point-by-point method is presented that allows the analyst to determine either absolute or relative curvatures and twists of thin plates. A full-field method is presented that will determine the absolute curvatures and twists of thin plates. Experiments have been performed that indicate that the methods do provide a simple and accurate means for determining the curvatures and twists of thin plates.

Introduction

Optical methods have been developed extensively in recent years to allow the analyst to determine the stresses that exist in a loaded, deformable body. For example, in the theory of small deflection of thin elastic plates and beams subjected to flexure, the stresses are directly related to the curvatures, namely $\partial^2 w/\partial x^2$, $\partial^2 w/\partial y^2$, and $\partial^2 w/\partial x \partial y$. The mixed second derivative $\partial^2 w/\partial x \partial y$ is sometimes referred to as the twist. Most experimental methods that are used to solve plate bending problems involve obtaining the deflections or the slopes and then performing differentiation(s) to yield the necessary curvatures for stress calculations [1–5]. However, either numerical or graphical differentiation will introduce errors that often reduce the accuracy of the experimental method. A more desirable approach is to determine curvatures directly from the experiment, in which case the differentiation procedure can be eliminated.

There are a few techniques currently available that will display curvature contours directly. A modified Lightenberg's reflection moiré method was developed by Heise [6]. Holographic interferometry was used by Toyooka [7] to determine contour maps of slope and curvature. Assa, Betzer, and Politch [8, 9] presented several methods to obtain slopes and curvatures by using single and double-frequency grating shearing interferometers. Chiang and Kao [10] also used a grating, in a manner similar to one of Assa's techniques, to obtain slope and curvature values for plates. Pointwise methods have been proposed by Duncan and Sabin [11], Chiang and Bailangadi [12], and Stetson [13]. It is noted that Stetson used moiré gratings to obtain curvatures at points on the plate surface.

In the present paper a shearing interferometric method is employed to measure curvature values of a specularly reflecting, flexed plate model. Two different techniques are proposed. One is a point-by-point scheme in which the curvature value is obtained from a single exposure of the film. The second is a full-field method, which yields the curvature contours of a flexed plate. The interpretations of the interference fringes are presented. Experiments are performed for each case to verify its feasiblity with emphasis on the use of Wollaston prism for shearing interferometry.

Optical System

The basic optical system is shown in Fig. 1. A planar object with a mirrorlike surface is illuminated by a plane wave from a laser source. The out-of-plane displacements of an object are most easily obtained by using illumination and a viewing system both of which are normal to the object surface. The wavefront reflected from the plate model passes through the lens system and is sheared into two wavefronts. The two wavefronts impinge on the photographic film, and the information of interest is recorded. If a Wollaston prism [14–18] with a 45 deg polarizer is used as the shearing mechanism in the setup shown in Fig. 1, then the image plane lateral shear, δ_1, is given by

$$\delta_1 = \alpha d' \tag{1}$$

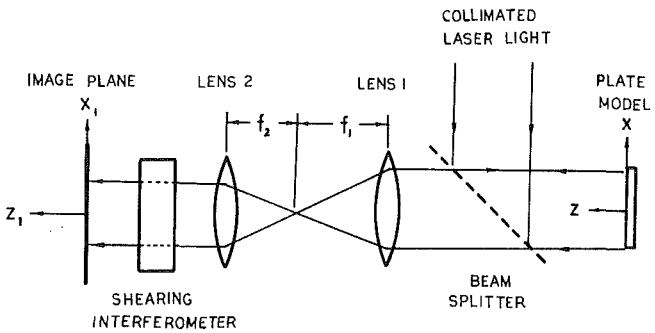


Fig. 1 A schematic view of optical system

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March, 1982; final revision, June, 1982. Paper No. 82-WA/APM-27.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where d' is the distance from the prism to the image plane and α is the prism diverging angle, given by

$$\alpha = 2(n_e - n_o) \tan \tilde{\theta} \tag{2}$$

with n_e and n_o the extraordinary and ordinary indices of refraction for the prism and $\tilde{\theta}$ the wedge angle for the prism.

Analysis—General

Let the object plane and image plane coordinates be given by (x,y,z) and (x_1,y_1,z_1), respectively, as shown in Fig. 1. Also, let $\mathbf{k}$ be the wave propagation vector where $|\mathbf{k}| = 2\pi/\lambda$ with λ defined as the wavelength of the illuminating laser light. Finally, let $\mathbf{r}$ be the vector position of the point of interest with respect to the origin of the (x,y,z) coordinate system. Then, if one suppresses the time dependence, the complex amplitude immediately after the wave is reflected by the object can be written

$$U(x,y) = a \exp [i\, \mathbf{k} \cdot \mathbf{r}]$$

or

$$U(x,y) = a \exp \left[i \frac{2\pi}{\lambda} z(x,y) \right] \tag{3}$$

where a is a constant representing the amplitude of the wavefront, $i = \sqrt{-1}$, and $z(x,y)$ is the phase factor for the reflected wave. For an undeformed planar plate model,

$$z(x,y) = z_0 \tag{4a}$$

with z_0 a constant at a fixed z location. If the object is deformed and the small deflection theory of thin plates is assumed valid, then $z(x,y)$ can be related to the deformation of the object surface by

$$z(x,y) = z_0 + 2w(x,y) \tag{4b}$$

where $w(x,y)$ is the displacement of the model surface in the z-direction. It is noted that if off-axis illumination is used in Fig. 1, a factor of $\cos \tilde{\theta}$ must be incorporated into (4*b*), $\tilde{\theta}$ being the angle between the z-axis and the illumination direction. Therefore, if the shear is given in the x-direction, then the complex amplitude distribution in the image plane (x_1,y_1) can be written as

$$U(-x_1,-y_1) = b \exp\left[i \frac{2\pi}{\lambda} \left\{ z\left(\frac{x_1 + \delta_1}{M}, \frac{y_1}{M} \right) + \frac{x_1 \alpha}{2} \right\} \right] + b \exp\left[i \frac{2\pi}{\lambda} \left\{ z\left(\frac{x_1}{M}, \frac{y_1}{M} \right) - \frac{x_1 \alpha}{2} \right\} \right] \tag{5}$$

where M is the lateral magnification factor of the imaging lens system, b is a constant, and α is the prism divergence angle. For simplicity, a common and constant-phase factor related to the light propagation through the prism has been omitted in equation (5). Furthermore, it is assumed that the two sheared wavefronts have equal amplitude b. The intensity distribution is the product of the complex field with its complex conjugate and is given by

$$I(-x_1,-y_1) = U(-x_1,-y_1) \cdot U^*(-x_1,-y_1)$$

where * denotes the complex conjugate. Therefore, the intensity can be written as

$$I(-x_1,-y_1) = 2b^2 \left\{ 1 + \cos\left[\frac{2\pi}{\lambda} \left(z\left(\frac{x_1 + \delta_1}{M}, \frac{y_1}{M} \right) - z\left(\frac{x_1}{M}, \frac{y_1}{M} \right) + \alpha x_1 \right) \right] \right\} \tag{6}$$

Equation (6) forms the basis for the various experimental procedures that have been developed for this work. Each scheme relies on the presence of the phase difference terms $z(x + \Delta x,y) - z(x,y)$ and the embedded grid term αx_1 to obtain the information of interest.

Analysis—Point-by-Point Method

If a photographic medium is exposed to the light reflected from an undeformed object, then the intensity in the image plane can be directly obtained by substitution of equation (4*a*) into (6) to give

$$I_1(-x_1,-y_1) = 2b^2 \left(1 + \cos\left(\frac{2\pi}{\lambda} \alpha x_1 \right) \right) \tag{7}$$

This intensity distribution represents a pattern of parallel periodic fringes with a sinusoidal variation in intensity along the x_1-direction. The fringes have a uniform pitch, p, given by $p = \lambda/\alpha$. This "grid" will be used as a "reference state" in the point-by-point method.

If a photographic medium is exposed to the light wave reflected from a deformed object surface, then the intensity in the image alone can be obtained by substitution of equation (4*b*) into (6). In addition, the shift δ_1 is assumed small enough so that terms of $0(\delta_1)$ are negligible in a Taylor's series expansion of $z(x_1 + \delta_1/M,\ y_1/M)$. Therefore, the intensity distribution in the image plane is given as

$$I_2(-x_1,-y_1) = 2b^2 \left[1 + \cos\left(\frac{2\pi}{\lambda} \left(2\delta_1 \frac{\partial w}{\partial x_1} + \alpha x_1 \right) \right) \right] \tag{8}$$

Equation (8) can be considered to be a diffraction grating of variable pitch whose spatial frequency is modulated by the term proportional to $\partial w/\partial x_1$. From the intensity patterns defined by equation (8) bright fringes occur when

$$2\delta_1 \frac{\partial w}{\partial x_1} + x_1 \alpha = N\lambda, \quad |N| = 0, 1, 2, \ldots \tag{9}$$

The partial derivative of N with respect to x_1 gives the spatial frequency component of the fringes at (x_1,y_1) along the x_1-direction and is written

$$\frac{\partial N}{\partial x_1} = \frac{\alpha}{\lambda} + \frac{2\delta_1}{\lambda} \frac{\partial^2 w}{\partial {x_1}^2} \tag{10}$$

If the concept of local pitch, $p(x,y)$, is introduced, we may regard this quantity as a continuous variable in the (x,y) plane. The reciprocal of equation (10) is defined as the local pitch $p'(x_1,y_1)$ in the x_1-direction. It is given by

$$p'(x_1,y_1) = \frac{\lambda}{\alpha + 2\delta_1 \dfrac{\partial^2 w}{\partial {x_1}^2}} \tag{11}$$

As a simpler example, if a similar analysis is used on (7), then the local spatial frequency of the fringes would be $\alpha/\lambda = 1/p$. In this case, the local spatial frequency equals the global spatial frequency of the carrier wave.

Equation (11) can be written as

$$\frac{\partial^2 w}{\partial {x_1}^2} \left(\frac{x_1}{M}, \frac{y_1}{M} \right) = \frac{\lambda}{2\delta_1} \left(\frac{1}{p'} - \frac{1}{p} \right)$$

or

$$\frac{\partial^2 w}{\partial x^2} (-x,-y) = \frac{M^2 \lambda}{2\delta_1} \left(\frac{1}{p'} - \frac{1}{p} \right) \tag{12}$$

where $p = \lambda/\alpha$, given in equation (7). Equation (12) demonstrates that changes in the grid pitch parallel to the shear direction are directly proportional to the object curvature in the shear direction. As the plate model is bent, the twist will also have an effect on the grid pattern corresponding to the undeformed plate state. This can readily be shown by finding the spatial frequency component of the distorted

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

fringes perpendicular to the shear direction. If one takes the partial derivatives of N with respect to y_1 and assumes continuity in the second derivatives, then the local pitch at (x_1, y_1) along y_1-direction is shown to be

$$p''(x_1, y_1) = \frac{\lambda}{2\delta_1 \left(\frac{\partial^2 w}{\partial x_1 \partial y_1}\right)} \tag{13}$$

Thus the twist is given as

$$\frac{\partial^2 w}{\partial x \partial y}(-x, -y) = \frac{M^2 \lambda}{2\delta_1 p''} = \frac{M^2 p \alpha}{2\delta_1 p''} \tag{14}$$

In this derivation, the curvature that leads to $p' < p$ is regarded as positive whereas if the curvature leads to $p' > p$, it is considered to be negative. Thus if one determines the pitch of the grid at position (x,y) both before and after the loading of the plate model, then the curvature and twist values at this point can be calculated from equations (12) and (14). It is noted that to increase the accuracy of this scheme the term $M^2\lambda/2\delta_1$ in equations (12) and (14) should be made as small as possible.

Analysis—Full-Field Method

In this method, the principle of obtaining moiré fringes due to the overlap of two shifted, deformed, high-frequency gratings is employed. Full-field contours related to the curvatures or twists are obtained using this concept.

The object is deformed prior to being photographed for the first exposure. The film is then translated or shifted a small amount prior to the second exposure. A second exposure of the deformed object is then recorded on the shifted photographic medium. The light intensity recorded by the first exposure is given in equation (8). If the film translation is along the x_1-direction and equal to δ_2, the incident light intensity for the second exposure can be written as

$$I_3(-x_1, -y_1) = 2b^2 \left\{1 + \cos\left[\frac{2\pi}{\lambda}\left(2\delta_1 \frac{\partial w}{\partial x_1}(x_1 + \delta_2, y_1) + (x_1 + \delta_2)\alpha\right)\right]\right\} \tag{15}$$

If it is assumed that δ_2 is small so that only the first two terms in the Taylor's series expansion of $\partial w/\partial x_1\ (x_1 + \delta_2, y_1)$ are obtained, then the expression for the total exposure is the sum of equations (8) and (15), which can be written as

$$I_t(-x_1, -y_1) = I_2(-x_1, -y_1) + I_3(-x_1, -y_1) = 4b^2 \left\{1 + \cos\frac{2\pi}{\lambda}\left(\delta_1 \delta_2 \frac{\partial^2 w}{\partial x_1^2} + \frac{\alpha \delta_2}{2}\right) \cdot \cos\frac{2\pi}{\lambda}\left[2\delta_1 \frac{\partial w}{\partial x_1} + \delta_1 \delta_2 \frac{\partial^2 w}{\partial x_1^2} + x_1 \alpha + \frac{\alpha}{2}\delta_2\right]\right\} \tag{16}$$

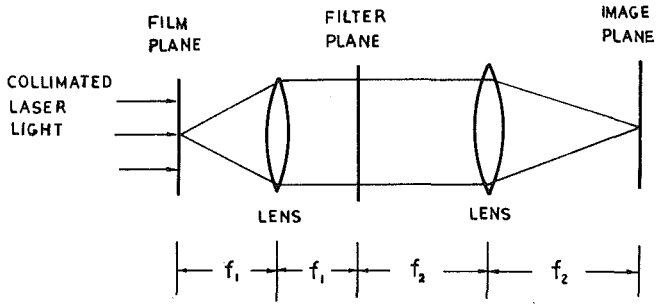


Fig. 2 The schematic of a basic Fourier filtering system

If the second term in equation (16) varies much faster than the first cosine term due to αx_1 (i.e., $|\partial^3 w/\partial x^3| << |\alpha/\delta_1\delta_2 + 2/\delta_2\ \partial^2 w/\partial x_1^2|$), then the expression given in (16) represents a high-frequency carrier wave cos $(2\pi/\lambda\ (x_1\alpha + \alpha\delta_2/2 + 2\delta_1\ \partial w/\partial x_1 + \delta_1\delta_2\ \partial^2 w/\partial x_1^2))$ amplitude modulated by the low-frequency term cos $(2\pi/\lambda\ (\delta_1\delta_2\ \partial^2 w/\partial x_1^2 + \alpha\delta_2/2))$. The high-frequency term is nulled when the first cosine term is equal to zero. In terms of the object plane coordinates, we can write

$$\frac{\partial^2 w}{\partial x^2}(-x, -y) = \frac{M^2\lambda}{2\delta_1\delta_2}\left(N - \frac{\delta_2}{p}\right), \quad |N| = 1/2, 3/2, \ldots \tag{17}$$

where $p = \lambda/\kappa$. Hence, if δ_2 is chosen arbitrarily to be $\delta_2 = (N' + 1/2)p$ where N' is an integer, then equation (17) becomes

$$\frac{\partial^2 w}{\partial x^2} = \frac{M^2 \lambda m}{2\delta_1\delta_2}, \quad |m| = 0, 1, 2, \ldots \tag{18}$$

where $m = N - (N' + 1/2)$ is the apparent fringe number.

If better fringe contrast is required, the transparency can be placed into a Fourier filtering arrangement as shown in Fig. 2. The Fourier filtering will block both the D.C. term and the high-frequency term in (16) so that the contrast of the filtered fringe pattern is enhanced. It can be shown that the filtered intensity distribution in this case is of the form [14]

$$I_f(-x_1, -y_1) = e\left[1 + \cos\frac{2\pi}{\lambda}\left(2\delta_1\delta_2 \frac{\partial^2 w}{\partial x_1^2} + \alpha\delta_2\right)\right] = 2e\cos^2\left[\frac{2\pi}{\lambda}\left(\delta_1\delta_2 \frac{\partial^2 w}{\partial x_1^2} + \frac{\alpha\delta 2}{2}\right)\right] \tag{19}$$

with e being a constant. Equation (19) yields the same fringe pattern as described by (16) and (18). However, (19) indicates that there is no background term and the zeros of the intensity pattern are proportional to the fringe of the cosine term. Therefore, the pattern will have enhanced fringe contrast.

In a manner analogous to equation (16), the mixed partial derivaties terms $\partial^2 w/\partial x_1 \partial y_1$, often designated as the twist, can be obtained. If the first shear, δ_1, is in the x_1-direction and if the second shear, δ_2, is in the y_1-direction, (15) becomes

$$I_3(-x_1, -y_1) = 2b^2\left\{1 + \cos\left[\frac{2\pi}{\lambda}\left(2\delta_1 \frac{\partial w}{\partial x_1}(x_1, y_1 + \delta_2) + \alpha x_1\right)\right]\right\} \tag{20}$$

The sum of equations (8) and (20) can then be written as

$$I_t(-x_1, -y_1) = 4b^2\left\{1 + \cos\left[\frac{2\pi}{\lambda}\left(\delta_1\delta_2 \frac{\partial^2 w}{\partial x_1 \partial y_1}\right)\right] \cdot \cos\left[\frac{2\pi}{\lambda}\left(2\delta_1 \frac{\partial w}{\partial x_1} + \delta_1\delta_2 \frac{\partial^2 w}{\partial x_1 \partial y_1} + \alpha x_1\right)\right]\right\} \tag{21}$$

If one uses the same arguments as described previously, the fringe pattern can be obtained either directly from the double-exposed film or by use of a Fourier filtering technique. Dark fringes occur when the first cosine term in equation (21) is zero, or

$$\frac{\partial^2 w}{\partial x \partial y} = \frac{M^2\lambda}{2\delta_1\delta_2} N, \quad |N| = 1/2, 3/2, 5/2, \ldots \tag{22}$$

Equations (18) and (22) give the interpretation formulas for contours of constant curvature $\partial^2 w/\partial x^2$ and contours of constant twist $\partial^2 w/\partial x \partial y$, respectively. Should the two shears in (8) and (15) be in the y_1-direction, a similar analysis will show that the fringes formed in this case describe curvatures along y-direction, i.e., $\partial^2 w/\partial y^2$. Thus a complete set of second derivatives, $\partial^2 w/\partial x^2$, $\partial^2 w/\partial y^2$, and $\partial^2 w/\partial x \partial y$ can be obtained.

An alternative procedure to this scheme is that instead of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

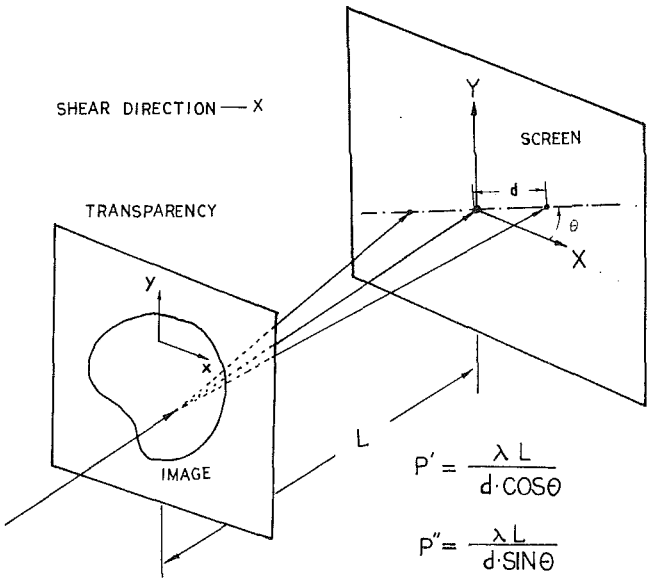


Fig. 3 Measurement of local pitches

photographing the loaded specimen twice, a single exposure of the film is made. Then a duplicate of the processed film is superimposed on the original film with a small shift δ_2 in the x_1-direction. The intensity expression is described by multiplication of equations (8) and (15) which yields

$$I_t'(-x_1,-y_1)=I_2\times I_3=\left[\cos\frac{2\pi}{\lambda}\left(\delta_1\delta_2\frac{\partial^2 w}{\partial x_1{}^2}+\frac{\partial 2\alpha}{2}\right)\right.$$

$$\left.+\cos\frac{2\pi}{\lambda}\left(2\delta_1\frac{\partial w}{\partial x_1}+\delta_1\delta_2\frac{\partial^2 w}{\partial x_1{}^2}+x_1\alpha+\frac{\alpha\delta_2}{2}\right)\right]^2 \quad (23)$$

It can be shown that the preceding equation leads to the same fringe pattern as described by equation (18). It is noted that fringe contours given by equation (22) can also be obtained by this alternative procedure if δ_2 is given in the y_1-direction.

Experimental Procedure

A schematic of the optical arrangement used in the experiments is given in Fig. 1. A Wollaston prism made of quartz with a divergence angle of approximately 1 deg is employed in this work as the sharing component. A 15 Mw Ne-Ne laser was used to illuminate the object. Afga-Gaevert 10E75AH film was used to record the patterns, with exposure times of ~0.5 sec. The experimental setup was mounted on an isolation table to eliminate extraneous vibrations. The objects used for the experimental work were a cantiliever plate (50 mm × 44.5 mm × 2.3 mm), a circular plate (31.6 mm × 2.6 mm), and a cantilever beam (31 mm × 9 mm × 3.18 mm). The surfaces of a typical photoelastic sheet obtained from Photoelastic, Inc. are quite flat. These untreated surfaces were suitable for use as the speculary reflecting surfaces in this work. Ultraflat black paint was sprayed on the rear surface of the model to reduce the unwanted rear surface reflections.

Experimental Work—Point-by-Point Methods

A cantilever beam is employed in this experimental investigation to verify the analytical results. The primary shear, δ_1, is equal to 0.6 mm and is in the x-direction along the beam axis. The out-of-plane free end deflection of the beam is 0.51 mm. For this special case, the curvature is linear along the beam axis. The local pitch of the embedded grid was measured by two different methods. In the first method, the photographic transparency was placed under a microscope

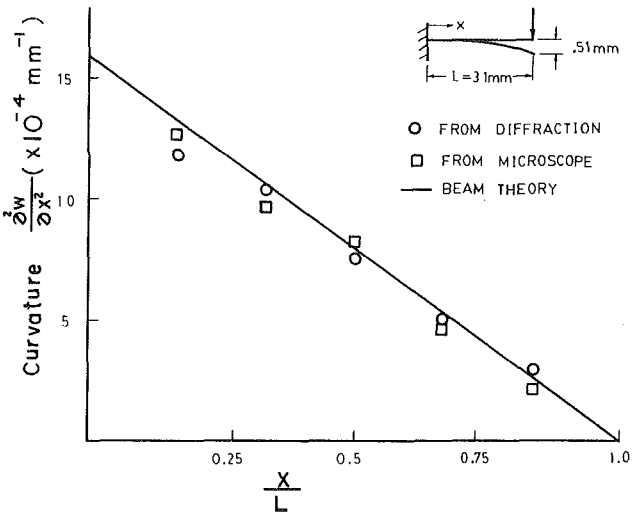


Fig. 4 Curvature of a cantilever beam

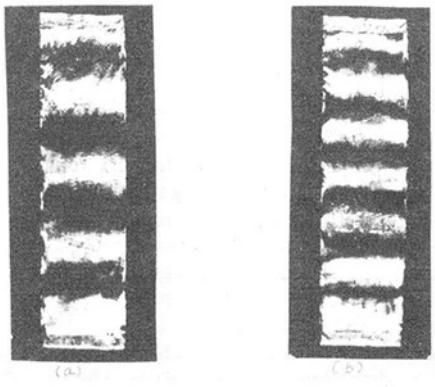


Fig. 5 $\partial^2 w/\partial x^2$ contours of the centilever beam with shears; (a) δ = 0.51 mm; (b) δ = 0.76 mm; and end deflection δ

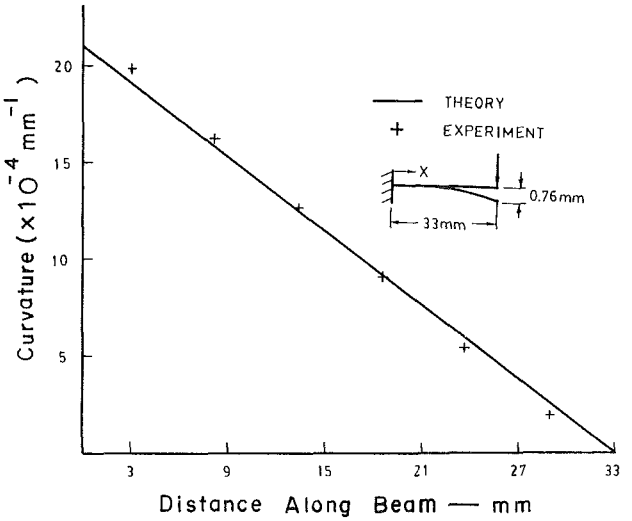


Fig. 6 $\partial^2 w/\partial x^2$ contour of a centilever beam; a comparison of theory and experimental results

and the local pitch was read directly from the fine scale in the viewfinder of the microscope. The second method utilized the diffraction property of the embedded grid. The developed film was placed in the path of a narrow unexpanded beam of laser light such that the point of interest is under the illumination. This concept is sketched in Fig. 3. Since the intensity distribution on the transparency has a periodic

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

character with a high frequency, then the diffracted light on a screen at a distance L away from the transparency will appear as three bright spots which correspond to 0, +1, and −1 diffraction orders. If one measures the components of the separation between any two neighboring diffraction orders in the x and y-directions, respectively, then the expression for p' and p'' at the point of interest are given by

$$p' = \frac{\lambda L}{d\cos\theta}, \quad p'' = \frac{\lambda L}{d\sin\theta} \tag{24}$$

where θ, d, and L are as shown in Fig. 3. It is noted that θ is measured from the $+x$ axis and $|\theta| < 90$ deg. The sign of the twist is determined for this case using the sign of sin θ. The results of the experimental work are presented in Fig. 4. Good arguments between the theory and experiment is obtained except near the fixed end of the cantilever beam. Errors in this region were found to be a result of variations in the support fixity and the initial curvatures introduced in this region by the support fixity.

Full Field Method

Three models are used for the experimental work. The cantilever beam model was subjected to out-of-plane, free end deflections of 0.51 mm and 0.76 mm. The primary and secondary shears for both cases is 0.62 mm and 1.41 mm, respectively, with both directed along the beam axis. The curvature fringes for these cases are shown in Fig. 5. Numerical values from Fig. 5(*b*) are plotted in Fig. 6 for comparison to the known theoretical results. Good agreement between the theoretical and experimental values is obtained.

A clamped circular plate was subjected to a centrally located out-of-plane deflections of 0.05 mm. The primary and secondary shears for this case are $\delta_1 = 0.62$ mm and $\delta_2 = 1.4$ mm, respectively. The curvature fringes for this case are shown in Fig. 7(*a*).

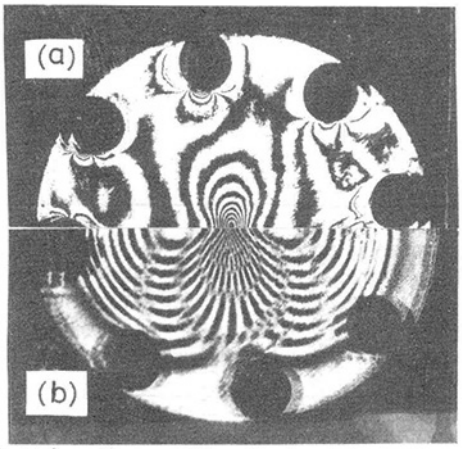


Fig. 7 $\partial^2 w/\partial x^2$ contours of the circular plate by two different techniques

In addition, curvature contours for this same case were generated by the moiré method using two shifted slope, fringe patterns, as shown in Fig. 7(*b*). A good match of fringe positions is obtained.

A centilever plate is used to obtain both twist contours and curvature contours. The applied, out-of-plane deflection was located approximately at the center of the plate and was eqaul to 0.51 mm. The primary shear for both the curvature and twist measurements was $\delta_1 = 0.62$ mm in the direction perpendicular to the fixed end of the support (y-direction). The secondary shears were equal to 2.82 mm, 1.41 mm, and 2.82 mm for Figs. 8(*a*), 8(*b*), and 8(*c*), respectively. It is noted that the number of fringes in 8(*c*) was approximately double the number of fringes displayed in 8(*b*). This result is correctly predicted by equation (22).

Discussion

It is noted that the two proposed methods can lead to different results, with the initial plate curvature introducing the distinguishing factor. If the point-by-point method is employed, the relative curvatures can be obtained if the grid line pitch p (for the image of the unloaded specimen) is considered as a function of position (x,y). If this $p(x,y)$ is compared with $p'(x,y)$ and $p''(x,y)$ in (12) and (14), then the calculated curvature and twist will yield the relative values. If the full-field method is employed then the loading is applied before taking exposures, and, therefore, both the curvatures due to the loading and the initial curvature will contribute to the resultant fringe pattern indicated in (19). As a consequence the absolute curvature fringes are established by use of the full-field methods. This can be seen from Fig. 7(*a*). The irregular fringes around the screws near the edge of the plate demonstrate that the absolute curvature has been present. These irregular fringes are due to the localized squeezing of the plate surface by the tightened screws. On the other hand, the moiré fringes in Fig. 7(*b*) give relative curvature fringes as usual. If the model surface is initially flat then the initial curvature is nonexistent and the two measurement techniques will yield the same results in general.

It is also noted that the derivations that were used in equations (8), (12), (14), (16), and (22) were deduced using finite shear. Since a finite amount of shear is introduced in the methods, the measured values of the curvatures are the values averaged over the distance of the shear. In the experiments, the measured values were considered to occur at the midpoint of the shear distance. The errors induced by a finite shear were investigated in detail by Boone [5]. Additional errors may be due to the measuring of physical qantities such as the amount of shear used in the experiment. The deviation of the experimental results from the plate theory was found to be less than 8 percent in this work.

It should be emphasized that the proposed methods

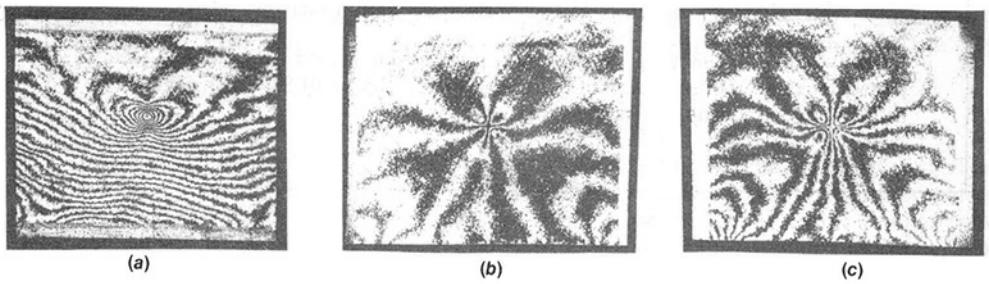


Fig. 8 Curvature and twist contours for the centilever plate with fixed end at the bottom of the plate. (a) $\partial^2 w/\partial y^2$; (b) $\partial^2 w/\partial x\partial y$; (c) $\partial^2 w/\partial x\partial y$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

determine curvatures directly rather than deflections or slopes. Thus the errors associated with differentiation process are eliminated.

Conclusion

Two shearing interferometric techniques are presented to obtain both the curvature and twist of a flexed plate. The point-by-point method provides a technique that will determine either the absolute or relative curvature and twist values at individual object points. The full-field method determines the absolute curvature and twist values at all points on the object surface. A Wollaston prism was employed as the shearing mechanism to demonstrate its effectiveness. Both analytical and experimental work are presented. The comparisons of analytical predictions to the experimental data have demonstrated that both techniques are viable experimental methods that may be used to determine the deformation of thin plates.

Acknowledgment

This work was completed in the Theoretical and Applied Mechanics Department at the University of Illinois. The authors are deeply appreciative for the financial support of the department and the moral support of its Chairman, Professor R. T. Shield.

References

1 Ligtenberg, F. K., "The Moiré Method—A New Experimental Method for the Determination of Moments in Small Slab Models," *Proc. SESA*, XII, 1954, pp. 83-98.

2 Theocaris, P. S., "Moiré Method in Plates," in: *Proc. Int. Assoc. Shell Struct. Symposium*, Warsaw (North Holland, Amsterdam), 1963, pp. 811-889.

3 Theocaris, P. S., "Moiré Patterns of Slope Contours in Flexed Plates," *Exp. Mech.*, Vol. 6, No. 4, 1966, pp. 212-217.

4 Hovanesian, J. D., and Varner, J., "Method for Determining the Bending Moments in Normal Loaded Thin Plates by Holographic Interferometry," in *Engineering Uses of Holography*, Robertson, E. R., and Harvey, J. M., Eds., Cambridge University Press, London, 1970, pp. 173-185.

5 Boone, P. M., "Determination of Slope and Strain Contours by Double-Exposure Shearing Interferometry," *Exp. Mech.*, Vol. 15, No. 8, 1975, pp. 205-302.

6 Heise, V., "A Moiré Method for Measuring Plate Curvatures," *Exp. Mech.*, Vol. 7, No. 1, 1967, pp. 47-48.

7 Toyooka, S., "Contour Mapping of the First and the Second Derivatives of Plate Deflection by Using Modulated Diffraction Gratings Made by Double-Exposure Hollography," *Optica Acta*, Vol. 25, No. 10, 1978, pp. 991-1000.

8 Assa, A., Betser, A. A., and Politch, J., "Recordings Slope and Curvature Contours of Flexed Plates Using a Grating Shearing Interferometer," *Appl. Opt.*, Vol. 16, No. 9, 1977, pp. 2504-2513.

9 Assa, A., Politch, J., and Betser, A., "Slope and Curvature Measurement by a Double-Frequency-Grating Shearing Interferometer," *Exp. Mech.*, Vol. 19, No. 4, 1979, pp. 129-137.

10 Chiang, F. P., and Kao, T. Y., "An Optical Method of Generating Slope and Curvature Contours of Bent Plates," *Int. J. of Solids Structures*, Vol. 15, No. 3, 1979, pp. 251-260.

11 Duncan, J. P., and Sabin, P. G., "Determination of Curvatures in Flexed Elastic Plates by the Martinelli-Ronchi Techniques," *Exp. Mech.*, Vol. 3, No. 12, 1963, pp. 285-293.

12 Chaing, F. P., and Bailangadi, M., "A Method for Direct Determination of Small Curvatures," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, 1975, pp. 29-31.

13 Stetson, K., "Moire Method for Determining Bending Moments From Hologram Interferometry," *Optics Tech.*, Vol. 2, No. 2, 1970, pp. 80-84.

14 Chao, Y. J., "On the Use of Shearing Interferometry for Slope and Curvature Measurements of Thin Plates," Ph.D. Thesis, Dept. of Theoretical and Applied Mechanics, Univ. of Illinois, Urbana, 1981.

15 Francon, M., and Mallick S., *Polarization Interferometers—Applications in Microscopy and Macroscopy*, Wiley, New York, 1971.

16 Chevalerias, R., Latron, Y., and Véret, C., "Methods of Interferometry Applied to the Visualization of Flows in Wind Tunnels," *J. Opt. Soc. Amer.*, Vol. 47, No. 8, 1957, pp. 703-706.

17 Black, W. Z., and Carrr, W. W., "Application of a Differential Interferometer to the Measurement of Heat Transfer Coefficients," *Rev. Sci. Instrum.*, Vol. 42, No. 3, 1971, pp. 337-340.

18 Small, R. D., Sernas, V. A., and Page, R. H., "Single Beam Schlieren Interferometer Using a Wollaston Prism," *Appl. Opt.*, Vol. 11, No. 4, 1972, pp. 856-862.

19 Bryngdahl, O., "Applications of Shearing Interferometry," *Progress in Optics*, Wolf, E., Ed., Vol. 4, 1964, pp. 37-83.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stability and Vibrations of Compressed, Aeolotropic, Composite Cylindrical Shells

J. B. Greenberg
Senior Lecturer.

Y. Stavsky[1]
Gerard Swope Professor of Mechanics.
Mem. ASME

Department of Aeronautical Engineering,
Technion-Israel Institute of Technology,
Haifa, Israel

A general method of solution, based on a complex finite Fourier transform, is adopted for the stability and vibration analysis of compressed, aeolotropic, composite cylindrical shells. A major feature of the solution method is its ability to handle both uniform and nonuniform conditions that hold at the boundaries of finite-length cylindrical shells. For the various shells investigated, an optimum winding angle was found for which a maximum frequency response and highest critical buckling load is attainable. Similar optimization was also discovered to be possible by controlling both/either shell heterogeneity and/or fiber orientation.

1 Introduction

Fiber-reinforced composites are increasingly being incorporated into a variety of structures in aeronautical, mechanical, and marine engineering applications where a high ratio of stiffness to weight is of paramount importance. As a consequence, the stability analysis and vibrational response of thin, composite, shell-type structures is of major interest in their design. In this study, laminated, anisotropic, cylindrical shells of finite length are considered, in which each layer is permitted an arbitrary fiber orientation, as is common in filament winding technology. Indeed, buckling of such shells has been the subject of a number of works which have been reviewed, for example, Tennyson [1] and Bert [2]. The buckling problem of *orthotropic* composite shells under axial compression was studied by Stavsky and Friedland [3], while free vibrations of such shells were dealt with by Stavsky and Loewy [4]. Greenberg and Stavsky [5] examined the vibrational response of compressed shells using a Love-type theory. A closed-form solution for simply supported (*SS*3) shells was obtained; results indicated that shell orthotropy and heterogeneity can have a significant effect on the buckling load and vibrational response of the system.

For layered *aeolotropic* shells the inverse method has generally been utilized for calculating critical buckling loads (see, for example, Cheng and Ho [6], Wu [7], and Herakovich and Johnson [8]). The difficulty in solving the eigenvalue problem directly for anisotropic shells, even for the *SS*3 boundary conditions, stems from the occurrence of the shear coupling terms that prevent the appearance of the chessboard mode of vibration present in orthotropic or isotropic shells. However, Padovan [9] has given a direct method for obtaining the critical buckling loads of anisotropic cylinders subjected to nonuniform axial prestress. The prestress must satisfy Dirichlet's conditions for a Fourier series. However, results are presented only for a single-layered orthotropc shell with *SS*3 boundary conditions and an anisotropic infinitely long cylinder, both of which provide a considerably simplified analysis. A similar general direct method has recently been suggested by Greenberg and Stavsky [10] for the free vibrations analysis of aeolotropic, composite cylindrical shells, based on the use of the complex finite Fourier transform.

In the present work, a Love-type theory is used for the vibrational response of composite, aeolotropic cylindrical shells subjected to axial compression. The equations of motion are solved using the aforementioned general direct method, the scope of which is extended analytically to cover the problem of nonuniform boundary condtions. The latter situation can frequently be caused, for example, by an improperly glued joint over part of the boundary. The study of this problem by Hoff and Soong [11] for an *isotropic* shell indicated that a reduction of the critical stress by a factor of one-half can ensue from such faulty manufacturing. The analysis in the present study treats this serious practical problem as it arises in laminated, aeolotropic cylindrical shells.

2 Equations of Motion

Consider a finite circular cylindrical shell made of layers of anisotropic materials whose elastic properties are thickness dependent only. In Fig. 1 the nondimensional axial coordinate, ξ, the circumferential coordinate, θ, and the normal, ζ, to the reference surface (of radius a) are illustrated. The reference surface, at which axial compression is applied is $\zeta = h/2$. The principal axes of each layer of material need not be parallel to the axial and circumferential axes. The fiber winding angle is denoted by ϕ.

[1] During 1981, Visiting Professor, Department of Mechanics and Structures, University of California at Los Angeles, Los Angeles, Calif. 90024.

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 1981; final revision, May 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

2.1 Strain-Displacement and Stress-Strain Relations. If Love's first approximation shell theory is applied, the strain and displacements vary linearly with the thickness coordinate, ζ, and the relevant details are unchanged from the development of Greenberg and Stavsky [5]. A similar remark is applicable to the dynamic equations. The point of departure from reference [5] for aeolotropic (or generally orthotropic) laminates relates to the stress-strain relations. Hooke's law for a generally orthotropic shell is written in terms of the elastic stiffness moduli.

$$\begin{bmatrix} \tau_\xi \\ \tau_\theta \\ \tau_{\xi\theta} \end{bmatrix} = \begin{bmatrix} E_{\xi\xi} & E_{\xi\theta} & E_{\xi s} \\ & E_{\theta\theta} & E_{\theta s} \\ \text{Symm.} & & E_{ss} \end{bmatrix} \begin{bmatrix} \epsilon_\xi \\ \epsilon_\theta \\ \epsilon_{\xi\theta} \end{bmatrix} \tag{1}$$

where, in general, $E_{\xi s}$, $E_{\theta s}$ are nonzero because arbitrary fiber orientation is permitted. As a consequence of (1) the stress-strain relations for the shell are:

$$\begin{bmatrix} \mathbf{N} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \epsilon^0 \\ \kappa \end{bmatrix} \tag{2}$$

where the elements of the full symmetric elastic coefficients matrices **A**, **B**, **D** are defined by:

$$(A_{ij}, B_{ij}, D_{ij}) = \int_{-h_1}^{h_2} E_{ij}(1, \zeta, \zeta^2) d\zeta \quad (i, j = \xi, \theta, s) \tag{3}$$

and the superscript 0 of ϵ refers to reference surface ($\zeta = 0$) values. The stress resultants and couples are defined by Stavsky and Loewy [4].

3 Displacement Equations of Motion

The displacement equations of motion are derived by substituting equation (2) into the dynamic equations of motion [5]. Vibratory motion of the entire system is taken to be simple harmonic with the same period, so that the temporal and spatial vibrations may be separated in the usual fashion.

$$(u^0, v^0, w^0) = Re\{(U_1, U_2, U_3) e^{i\lambda t}\} \tag{4}$$

where *Re* denotes the real part of, and the U_i are functions of ξ and θ. Utilizing (4) in the equations of motion produces a linear eighth-order system of partial differential equations of the form:

$$\sum_{j=1}^{3} L_{ij} U_j = 0 \quad (i = 1,2,3) \tag{5}$$

The functional operators L_{ij} are given in the Appendix. For the case of a shell possessing regular orthotropy the preceding system reduces to the equations solved by Greenberg and Stavsky [5].

4 Solution of Eigenvalue Problem by Finite Complex Fourier Transform

To solve the equations of motion the general direct method described by Greenberg and Stavsky [10] is adopted. It should be emphasized that application of this method is predicted on the linearity of the governing displacement equations. Therefore, although Love-type theory is used here, the method can be just as readily utilized if any other linear theory (e.g., Flügge) were to be applied.

The essence of the method involves taking the finite complex Fourier transform of the governing equations (5) (with respect to θ) thus producing a transformed set of *ordinary* differential equations for the (now transformed)) displacements $U_j^{(n)}$:

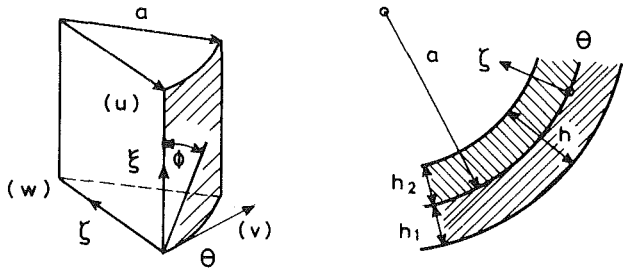


Fig. 1 Filament wound cylindrical shell notation

$$\sum_{j=1}^{3} L_{ij}^{(n)} U_j^{(n)}(\xi) = 0 \quad (j = 1,2,3;\ n = 0,1,2,\ \ldots) \tag{6}$$

where n characterizes a circumferential mode. The analytic solution of these equations can be shown to be:

$$U_j(\xi, \theta) = Re\left\{ \sum_{m=0}^{\infty} \frac{1}{2\pi} \left[\sum_{s=1}^{8} A_s^{(m)} d_j^{(m)} e^{\mu_s^{(m)} \xi} \right] e^{-im\theta} \right\} \tag{7}$$

where $d_j^{(m)}$ are certain differential operators and $\mu_s^{(m)}$ are the roots (assumed distinct) of an eight-order polynomial auxiliary equation. For a given value of n there are eight constants $A_s^{(n)}$ to be determined from the four boundary conditions at each end of the cylinder. Upon applying these conditions, the requirement of a nontrivial solution to the ensuing characteristic equations helps to furnish the lowest frequency of vibration of the shell (since in equation (7) the $\mu_s^{(m)}$ are implicitly functions of the frequency λ).

A major advantage of the use of this direct method for tackling the eigenvalue problem is its ability to readily handle *any* conditions that are applied at the boundaries of finite length shells. In particular the problem of *nonuniform* boundary conditions (caused, for example, by an improperly glued joint over part of the boundary) that was studied by Hoff and Soong [11] for an *isotropic* shell can be extended to the case of a laminated aeolotropic shell.

Let the nonuniform boundary conditions be described as follows:

$$\begin{aligned} {}^{(p)}G(\xi, \theta) = \quad & \sum_{j=1}^{3} {}^{(p)}\alpha_{kj} U_j = 0 \qquad -\pi < \theta < \theta_1^{(p)} \\ & \sum_{j=1}^{3} {}^{(p)}\beta_{kj} U_j = 0 \qquad \theta_1^{(p)} \le \theta \le \pi \\ & \qquad\qquad (k = 1,2,3,4;\ p = 0,1) \end{aligned} \tag{8}$$

where ${}^{(p)}\alpha_{kj}$, ${}^{(p)}\beta_{kj}$ are linear (possibly differential) operators, and the angle, $\theta_1^{(p)}$, associated with the nonuniformity may be different at each end of the cylinder. Transforming (8) and making use of equations (7) leads, after some algebra, to the necessity to solve an infinite-dimensional determinantal equation in order to determine the required eigenvalues:

$$\begin{bmatrix} \mathbf{C}^{(0,0)} & \mathbf{C}^{(1,0)} & \cdots \\ \mathbf{C}^{(1,0)} & \mathbf{C}^{(1,1)} & \cdots \\ \mathbf{C}^{(2,0)} & \mathbf{C}^{(2,1)} & \mathbf{C}^{(2,2)} \cdots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \mathbf{A}^{(0)} \\ \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = 0 \tag{9}$$

where $\mathbf{A}^{(n)}$ is an (8 × 1) vector whose elements are the eight constants of integration associated with mode n, and each 8

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

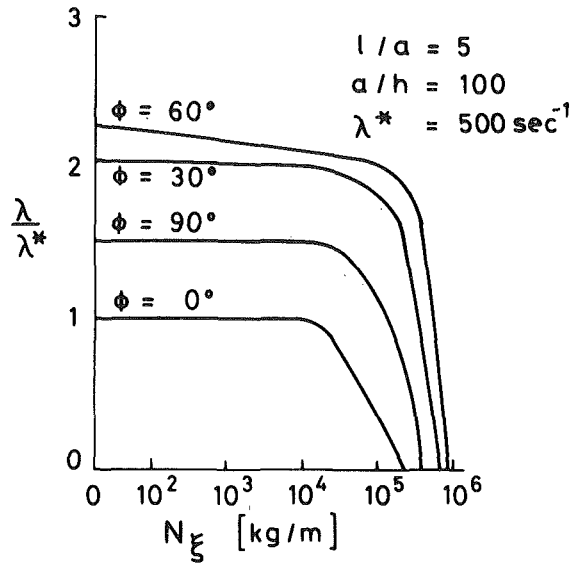


Fig. 2 Frequencies of clamped UHMG cylindrical shells with axial compression

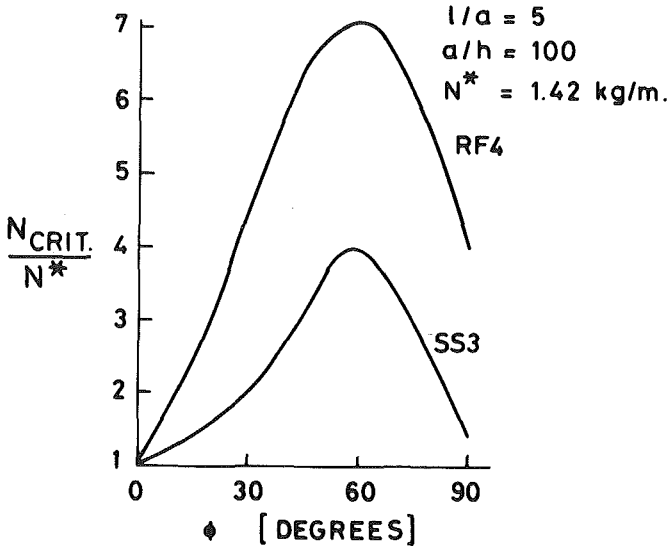


Fig. 3 Effect of winding angle on critical load of HSG cylindrical shells

$\times$ 8 submatrix $\mathbf{C}^{(n,m)}$ contains coefficients resulting from application of the eight boundary conditions (equation (8)). Note that for the case of *nonuniform* boundary conditions the off-diagonal elements $\mathbf{C}^{(n,m)}$ are *nonzero*, i.e., a coupling between the circumferential modes is implied. However, for *uniform* boundary conditions the coupling disappears enabling each mode n to be solved separately, i.e., $\mathbf{C}^{(n,m)} = 0$ $n \neq m$. In practice, of course, only a finite number of modes can be considered, and convergence of the solutions obtained using consecutively increasing finite numbers of modes must be established.

5 Approximate Solution of Transformed Equations

Despite the existence of the analytic solution to the displacement equations of motion it is often computationally more convenient to find an approximate solution using a finite-difference method. Locating the roots of the determinantal equation for λ can be a lengthy process if no order-of-magnitude estimate is available to initialize the search. In addition, during the search itself rather large or small real parts of the roots of the auxiliary equation for Φ frequently occur, inevitably causing computer over or underflow. Although the finite-difference method is approximate it bears the advantage of being free of the aforementioned difficulties and, in addition, yields many eigenvalues rapidly and directly. The transformed equations (8) are written in standard finite-difference form at discrete points ordered along the length of the cylinder (see, for example, Fox [12]). The differential equations are thus replaced by a set of linear algebraic equations, with the coupling of the circumferential modes being introduced through appropriate application of the boundary conditions. Nontriviality of the solution for the displacements requires that

$$\det(\hat{\mathbf{L}} - \lambda^2 \mathbf{I}) = 0 \tag{19}$$

where $\hat{\mathbf{L}}$ is a matrix of finite-difference weights and coefficients of $L_{ij}^{(n)}$ and $\mathbf{I}$ is the unit matrix. Formally the matrix $\hat{\mathbf{L}}$ is infinitely large. In practice, however, only a finite number of circumferential periods is considered. Once the elements of $\hat{\mathbf{L}}$ have been supplied standard computer library subroutines can be called for calculating the required eigenvalues. Finally the critical buckling load is found using Ziegler's [13] stability criterion. That value of N_ξ^0 for which the frequency vanishes corresponds to the required limiting load.

6 Calculated Results

To investigate the effects of material anisotropy on the frequency response and buckling of axially compressed, composite cylindrical shells the following materials were considered [14]:

(*a*) Ultra-High Modulus Graphic Epoxy (UHMG)

$E_{11} = 3 \times 10^{11}\,\mathrm{N/m^2}$; $E_{22} = 6.2 \times 10^{9}\,\mathrm{N/m^2}$;

$E_{66} = 4.1 \times 10^{9}\,\mathrm{N/m^2}$;

$\nu_{12} = 0.26$; $\rho = 1.6 \times 10^{3}\,\mathrm{kg/m^3}$.

(*b*) High-Strength Graphite Epoxy (HSG)

$E_{11} = 1.24 \times 10^{11}\,\mathrm{N/m^2}$; $E_{22} = 1.03 \times 10^{10}\,\mathrm{N/m^2}$;

$E_{66} = 5.5 \times 10^{9}\,\mathrm{N/m^2}$;

$\nu_{12} = 0.27$; $\rho = 1.5\,\mathrm{kg/m^3}$.

For a given fiber orientation the values of the transformed elastic coefficients were calculated from the preceding data using the formulas of Ambartsumyan [15]. For all the shells the total thickness and radius were assumed constant and equal to 0.4 and 40, respectively.

6.1 Uniform Boundary Conditions. Two possible sets of boundary conditions were specified: (see [16] for the notation):

*RF*4: Clamped $u = v = w = w_{,\xi} = 0$ at $\xi = 0, l/a$

*SS*3: Simply Supported $N_\xi = M_\xi = v = w = 0$ at $\xi = 0, l/a$

6.2 Single-Layered Shells. The frequency response of a clamped UHMG shell as a function of the (constant) applied load, for various fiber winding angles is illustrated in Fig. 2.

Using Ziegler's [13] dynamic stability criterion the critical buckling load values are determined from the points of intersection of the various curves with the N_ξ-axis. With regard to the frequency response, an applied load of less than 10^4 kg/m produces no radical change from the free vibrations frequency. The enhancement induced by winding the fibers at 60 deg, rather than leaving the shell orthotropic, involves a factor of about 2.3. For a higher load of, say, 10^5 kg/m this frequency enhancement factor increases to 4.1 and, in fact, finally the orthotropic shell buckles under a smaller load than the anisotropic one.

In Fig. 3 the influence of the fiber winding angle on the critical buckling load of an HSG shell is exhibited for both

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

clamped and simply supported boundary conditions. It is seen that an optimum winding angle of about 60 deg exists at which the critical load attains a maximum value. Define an increase factor, F, as the ratio of the maximum to minimum critical loads for the range of winding angles 0–90 deg. Then for the clamped shell this factor is 6.5 whereas for the simply supported shell it drops by about 40 percent to 3.95. This reduction is indicative of the important role the boundary conditions play in determining the critical load. It is interesting to note that the *greatest* change effected by altering the boundary conditions is for the *orthotropic* shell having $\phi = 90$ deg. The factor involved is:

$$N_{\text{crit}}^{RF4}/N_{\text{crit}}^{SS3} = 2.82 \tag{20}$$

The curves also serve to highlight the importance of precision in filamentary construction. For example, a clamped HSG shell having imperfect winding of $\pm$ 5 deg at 70 deg can produce a reduction or increase in the critical buckling load of between 10–15 percent.

In Fig. 4 the effect of shell thickness on the critical load is shown for an UHMG shell. Once again the characteristic maxima are observed—increase factors of 6.1, 11.3, and 2.5 correspond to the three a/h ratios 250, 375,500, respectively. Evidently, the thinner the shell the less pronounced does the effect of varying the winding angle become.

It may thus be concluded from the previously described results that for single-layered shells control of the critical buckling load (and frequency response) may be brought about by a suitable combination of fiber winding angle, boundary conditions, and shell geometry.

6.3 Bilayered Shells.

6.3.1 Effect of Fiber Orientation in a Single Material, Bilayered Shell. Consider a shell composed of two equally thick layers of a certain material, such that the orientation of the fiber of the inner layer is ϕ_1 while that of the outer layer is ϕ_2. The sum of these two angles is 90 deg.

Figure 5 shows how various combinations of ϕ_1 and ϕ_2 alter the critical buckling loads for an HSG shell. Note the two conspicuous maxima, and the single minimum at 45 deg (for the latter, the shell is effectively single-layered). The occurrence of the maxima and comparison with the curves corresponding to a single-layered shell wound completely at either ϕ_1 or ϕ_2 indicates that this sort of *bilayered* construction from a *single material* can lead to buckling loads that transcend those of a corresponding *single-layered* shell. Here a factor as high as 6.2 for angles $\phi_1 = 72$ deg, $\phi_2 = 18$ deg, is clearly observed.

6.3.2 Effect of Relative Layer Thickness in Bilayered Composite Shells. Let a two-layered shell be composed of an inner layer of UHMG and an outer layer of HSG. The relative thickness of the two layers will be permitted to vary between zero and one and h_1/h will be the relative thickness of UHMG. It is assumed that the fibers of the two materials are orientated at the same angle. In Fig. 6 the dependence of the critical buckling load on h_1/h is illustrated for a variety of values of the winding angle. Greenberg and Stavsky [10] have reported that the natural frequency of such shells can be optimized by an appropriate combination of shell heterogeneity and winding angle. A similar phenomenon is observed in Fig. 6. For the shell shown the best winding angle is 60 deg. It provides the highest critical buckling loads for all values of h_1/h and ϕ. Furthermore, for the 60 deg shell itself there is an optimal combination of material thicknesses, viz, $h_1/h = 0.45$ for which the increase factor is 25. A similar maximum is noted for the 30 deg shell, here $h_1/h = 0.39$ with the increase factor being 17. However, little improvement is brought about by changing the heterogeneity for the limiting orthotropic shells, although it is worth noting that the curves corresponding to these two shells intersect, so that the shell

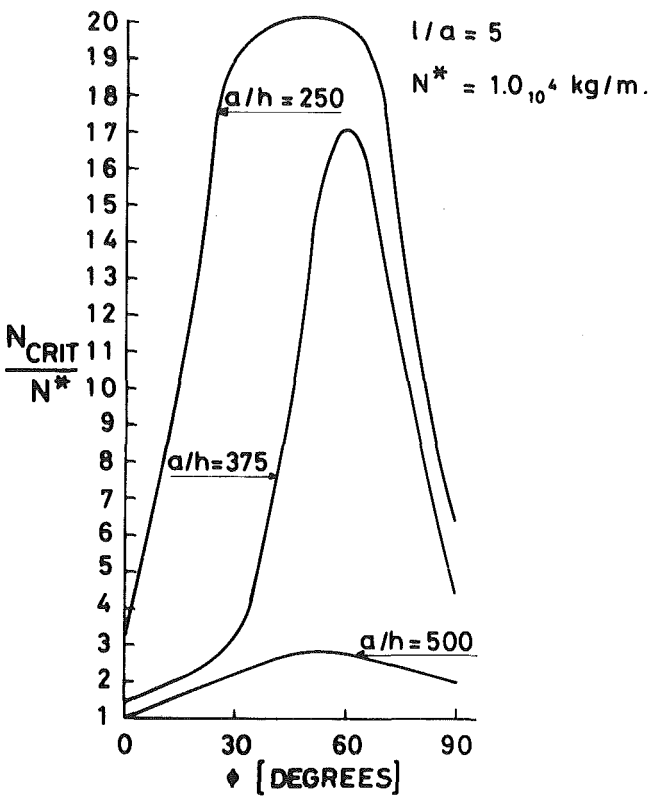


Fig. 4 Effect of shell thickness on critical load of clamped UHMG cylindrical shells

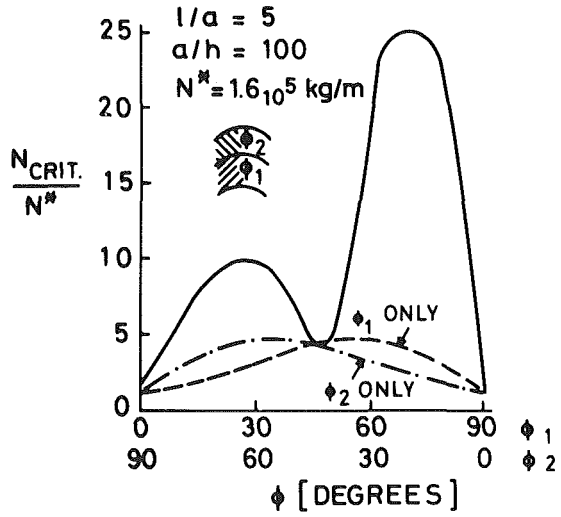


Fig. 5 Effect of winding angles on critical load of bilayered, simply supported, HSG cylindrical shells

having 35 percent UHMG has the same critical load for $\phi = 0$ or 90 deg.

6.3.3 Effect of Quasi-Orthotropic Assumption in Bilayered Shells. If a bilayered shell is composed of two equal layers of a single material, the inner layer being wound at $-\phi$ deg and the outer layer at $+\phi$ deg then the following elastic coefficients are identically zero: $B_{\xi\xi}, B_{\theta\theta}, B_{ss}, B_{\xi\theta}, A_{\xi s}, A_{\theta s}, D_{\xi s}, D_{\theta s}$. However, $B_{\xi s}$ and $B_{\theta s}$ are generally nonzero. A popular way of estimating the buckling loads of such ply-wound shells is to neglect $B_{\xi s}$ and $B_{\theta s}$, so that the stiffness coefficients matrix becomes similar in form to that of an orthotropic shell. For simply supported boundary conditions

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

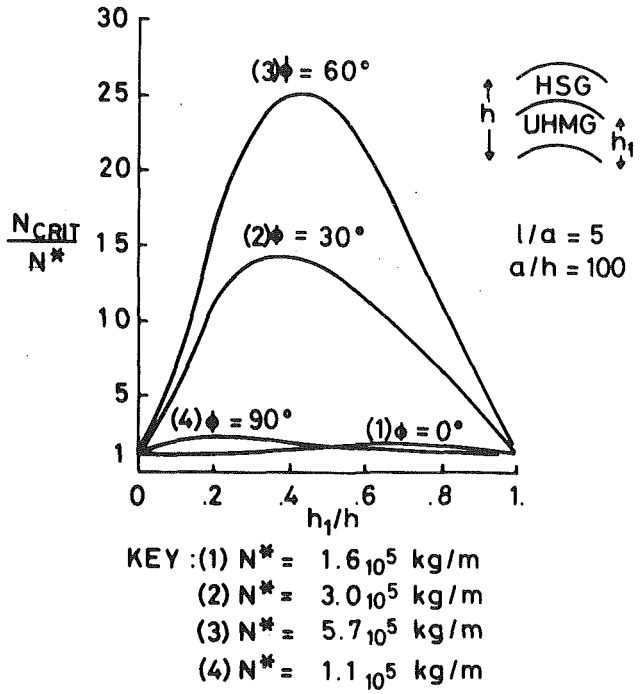


Fig. 6 Effect of relative thickness and winding angle on critical load of bilayered, UHMG/HSG simply supported, cylindrical shells

the buckling load can then be readily calculated (see, for example, Hirano [17]).

In the present work the validity of this approach is examined. Consider a ply-wound glass-epoxy shell. Relevant data is:

$$E_{11} = 7.8 \times 10^6 \text{psi}; E_{22} = 2.6 \times 10^6 \text{psi}$$

$$\nu_{12} = 0.25; E_{66} = 1.3 \times 10^6 \text{psi}$$

In Fig. 7 typical results obtained with and without the assumption of quasi-orthotropy are shown. Clearly the estimated critical load using the assumption does not give an accurate description of the behavior of the buckling load as a function of winding angle. The discrepancy is worst at about 30 deg where the approximate theory *overestimates* by a factor of 2. At 52 deg the theory is accurate but subsequently it underestimates the critical load, in the worst instance by about 1.7. It must be thus concluded that rule-of-thumb, quasi-orthotropic estimates of buckling loads for ply-wound shells should be used with extreme caution.

Acknowledgments

This paper was presented at the 15th Congress for Theoretical and Applied Mechanics in August 1980, in Toronto, Canada. This study was supported by Technion Research Funds and the Fund for the Advancement of Research. Y. Stavsky appreciates the helpful discussions of the paper with Professor N. J. Hoff of the Aeronautics and Astronautics Department of Stanford University, Professor J. Singer of the Aeronautical Engineering Department at Technion, and Professors S. B. Dong and L. A. Schmit, Jr., of the Mechanics and Structures Department at UCLA.

References

1 Tennyson, R. C., "Buckling of Laminated Composite Cylinders; A Review," *Composites*, 1975, pp. 17–24.

2 Bert, C. W., "Analysis of Shells," *Composite Materials*, Vol., Part 1, Chamis, C. C., Ed., Academic Press, New York, 1975, pp. 150–200.

3 Stavsky, Y., and Friedland, S., "Stability of Heterogeneous Orthotropic Cylindrical Shells in Axial Compression," *Israel Journal of Technology*, Vol. 7, 1969, pp. 111–119.

4 Stavsky, Y., and Loewy, R., "On Vibrations of Heterogeneous Orthotropic Cylindrical Shells," *Journal of Sound and Vibration*, Vol. 15, 1971, pp. 235–256.

5 Greenberg, J. B., and Stavsky, Y., "Buckling and Vibration of Orthotropic Composite Cylindrical Shells," *Acta Mechanica*, Vol. 36, 1980, pp. 15–29.

6 Cheng, S., and Ho, B. P. C., "Stability of Heterogeneous Aeolotropic Cylindrical Shells Under Combined Loading," *AIAA Journal*, Vol. 1, 1963, pp. 892–898.

7 Wu, C. H., "Buckling of Anisotropic Circular Cylindrical Shells," Ph.D. Thesis, Case Western Reserve University, Cleveland, Ohio, 1971.

8 Herakovich, C. T., and Johnson, E. R., "Buckling of Composite Cylinders Under Combined Compression and Torsion—Theoretical/Experimental Correlation," *ASTM Symposium on Test Methods and Design Allowables for Fibrons Composite*, Oct. 1969, Dearborn, Mich.

9 Padovan, J., "Frequency and Buckling Eigenvalues of Anisotropic Cylinders Subjected to Nonuniform Lateral Prestress," *International Journal of Solids and Structures*, Vol. 7, 1971, pp. 1449–1466.

10 Greenberg, J. B., and Stavsky, Y., "Vibration of Laminated Filament-Wound Cylindrical Shells," *AIAA Journal*, Vol. 19, 1981, pp. 1055–1062.

11 Hoff, N. J., and Soong, T. C., "Buckling of Axially Compressed Circular Cylindrical Shells With Non-Uniform Boundary Conditions," *Symposium on Thin Walled Steel Structures—Their Design and Use in Building*, University College of Swansea, Sept. 1967, pp. 61–80.

12 Fox, L., *The Numerical Solution of Two-Point Boundary Problems in Ordinary Differential Equations*, Oxford University Press, 1957.

13 Ziegler, H., "On the Concept of Elastic Stability," *Advances in Applied Mechanics*, Vol. 4, 1965, pp. 357–403.

14 Vicario, A. A., Sr., and Toland, R. H., "Failure Criteria and Analysis of Structural Components," *Composites Materials*, Vol. 7, Academic Press, New York, 1975, pp. 52–97.

15 Ambartsumyan, S. A., "Theory of Anisotropic Shells," NASA TT F-118, May, 1964.

16 Hoff, N. J., and Soong, T. C., "Buckling of Circular Cylindrical Shells in Axial Compression," *International Journal of Mechanical Sciences*, Vol. 7, 1965, pp. 489–520.

17 Hirano, Y., "Buckling of Angle-Ply Laminated Circular Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, 1979, pp. 233–234.

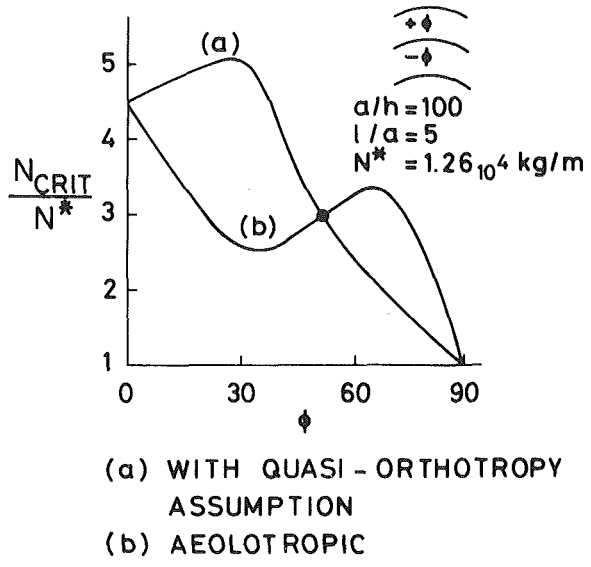


Fig. 7 Effect of quasi-orthotropy assumption on critical loads of bilayered, simply supported, glass-epoxy cylindrical shells

APPENDIX

The functional differential operators L_{ij} are given by

$$L_{11} = 2A_{s\xi}(\ )_{,\xi\theta} + A_{ss}(\ )_{,\theta\theta} + A_{\xi\xi}(\ )_{,\xi\xi} + \Omega^2(\ )$$

$$L_{12} = \left[A_{ss} + A_{\xi\theta} - \frac{1}{a}(B_{\xi\theta} + 2B_{ss})\right](\ )_{,\xi\theta}$$

$$+ \left(A_{s\theta} - \frac{B_{s\theta}}{a}\right)(\ )_{,\theta\theta} + \left(A_{\xi s} - \frac{2B_{\xi s}}{a}(\ )_{,\xi\xi}\right.$$

$$L_{13} = -\frac{B_{\xi\xi}}{a}(\ )_{,\xi\xi\xi} - \frac{B_{s\theta}}{a}(\ )_{,\theta\theta\theta} - \frac{3B_{s\xi}}{a}(\ )_{,\xi\xi\theta}$$

$$- \frac{1}{a}(2B_{ss} + B_{\xi\theta})(\ )_{,\xi\theta\theta} -$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-A_{s\theta}(\ )_{,\theta}-\left(A_{\xi\theta}+\frac{\Omega^2}{a}\frac{R_1}{R_0}\right)(\ )_{,\xi}$$

$$L_{21}=\left[A_{ss}+A_{\xi\theta}-\frac{1}{a}\left(B_{\xi\theta}+B_{ss}\right)\right](\ )_{,\xi\theta}$$

$$+\left(A_{s\theta}-\frac{B_{s\theta}}{a}\right)(\ )_{,\theta\theta}+\left(A_{s\xi}-\frac{B_{s\xi}}{a}\right)(\ )_{,\xi\xi}$$

$$L_{22}=\left(2A_{s\theta}-\frac{5B_{s\theta}}{a}+\frac{3D_{s\theta}}{a^2}\right)(\ )_{,\xi\theta}+\left(A_{\theta\theta}-\frac{2B_{\theta\theta}}{a}+\frac{D_{\theta\theta}}{a^2}\right)(\ )_{,\theta\theta}+$$

$$+\left(N_{\xi}^{0}+A_{ss}-\frac{3B_{ss}}{a}+\frac{2D_{ss}}{a^2}\right)(\ )_{,\xi\xi}$$

$$+\left(1-\frac{3R_1}{aR_0}+\frac{2R_2}{a^2R_0}\right)\Omega^2(\ )$$

$$L_{23}=\frac{1}{a}\left(-B_{s\xi}+\frac{D_{s\xi}}{a}\right)(\ )_{,\xi\xi\xi}+\frac{1}{a}\left(-B_{\theta\theta}+\frac{D_{\theta\theta}}{a}\right)(\ )_{,\theta\theta\theta}-$$

$$-\frac{1}{a}\left[B_{\theta\xi}+2B_{ss}-\frac{1}{a}\left(D_{\theta\xi}+2D_{ss}\right)\right](\ )_{,\xi\xi\theta}$$

$$+\frac{3}{a}\left(-B_{\theta s}+\frac{D_{\theta s}}{a}\right)(\ )_{,\xi\theta\theta}+$$

$$+\left[-A_{\theta\theta}+\frac{B_{\theta\theta}}{a}+\Omega^2\left(-\frac{R_1}{aR_0}+\frac{R_2}{a^2R_0}\right)\right](\ )_{,\theta}$$

$$+\left(-A_{s\theta}+\frac{B_{s\theta}}{a}\right)(\ )_{,\xi}$$

$$L_{31}=-L_{13}$$

$$L_{32}=\frac{1}{a}\left(B_{\xi s}-\frac{2D_{\xi s}}{a}(\ )_{,\xi\xi\xi}+\frac{1}{a}\left(B_{\theta\theta}-\frac{D_{\theta\theta}}{a}\right)(\ )_{,\theta\theta}+\right.$$

$$+\frac{1}{a}\left[2B_{ss}+B_{\xi\theta}-(4D_{ss}+D_{\xi\theta})/a\right](\ )_{,\xi\xi\theta}+$$

$$+\frac{1}{a}\left(3B_{s\theta}-\frac{4D_{s\theta}}{a}\right)(\ )_{,\xi\theta\theta}+\left(A_{\theta\theta}\right.$$

$$\left.-\frac{B_{\theta\theta}}{a}-\frac{2R_2}{a^2R_0}\Omega^2\right)(\ )_{,\theta}+\left(A_{\theta s}-\frac{2B_{\theta s}}{a}\right)(\ )_{,\xi}$$

$$L_{33}=-\frac{D_{\xi\xi}}{a^2}(\ )_{,\xi\xi\xi\xi}-\frac{D_{\theta\theta}}{a^2}(\ )_{,\theta\theta\theta\theta}-\frac{2}{a^2}(2D_{ss}+D_{\theta\xi})(\ )_{,\xi\xi\theta\theta}-$$

$$-\frac{4}{a^2}D_{\theta s}(\ )_{,\xi\theta\theta\theta}-\frac{4}{a^2}D_{s\xi}(\ ),\xi\xi\xi\theta-\frac{4B_{\theta s}}{a}(\ )_{,\xi\theta}$$

$$\left(-\frac{2B_{\theta\theta}}{a}-\frac{R_2}{a^2R_0}\Omega^2\right)(\ )_{,\theta\theta}+\left(N_{\xi}^{0}\frac{2B_{\theta\xi}}{a}\right.$$

$$\left.-\frac{R_2}{a^2R_0}\Omega^2\right)(\ )_{,\xi\xi}\left[(-A_{\theta\theta}+\Omega^2)(\ )\right.$$

where

$$\Omega^2=R_0(a\lambda)^2$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

S. L. Lau
Research Officer.
Mem. ASME

Y. K. Cheung
Professor and Head.

S. Y. Wu
Assistant Computer Officer.

Department of Civil Engineering,
University of Hong Kong,
Hong Kong

A Variable Parameter Incrementation Method for Dynamic Instability of Linear and Nonlinear Elastic Systems

A variable parameter incrementation method is proposed and then applied to the determination of parametric instability boundary of columns. Attention is particularly paid to the geometrically nonlinear problems including the instability of nonlinear vibrations. Although only beam and column problems are treated at present, the approach is believed to be general in methodology. This method is not subjected to the limitations of small exciting parameters and weak nonlinearity.

1 Introduction

In the computer analysis of nonlinear structures, the incremental method is in general regarded as the most suitable one among various procedures. However, the method has not received sufficient attention in all areas and its potentials remains to be explored in full.

In a recent paper [1], the authors have presented an amplitude incremental variational principle. Based on this principle an incremental time-space finite element method [2] has been developed and successfully applied to the nonlinear periodic structural vibration problems. It can be expected that a great deal of research work can be done along this line. Moreover, this kind of incremental method could also be extended to other related problems and even to other physical systems in which periodic solutions are sought.

In this paper, efforts are devoted to the development of the incremental method applicable to parametric instability problems. In particular, attention is drawn to geometrically nonlinear problems including the instability of nonlinear vibrations. Although only beams and columns are discussed at present, the approach is believed to be general in methodology. The main advantages of the proposed method are twofold. First, it is not subjected to limitations of small exciting parameters and weak nonlinearity. Second, it is particularly simple and economical for computer implementation because only linear algebraic equations have to be formed and solved in each incremental step.

2 Parametric Instability of Simple Column

To demonstrate the variable parameter incrementation method, consider first of all the determination of the instability boundary of a hinged-hinged column excited by axial harmonic force. The differential equation for this problem is known to be

$$\mathrm{E}I\frac{\partial^4 w}{\partial x^4} - P\cos 2\omega t\frac{\partial^2 w}{\partial x^2} + \rho A\frac{\partial^2 w}{\partial t^2} = 0 \qquad (2.1)$$

where E = Young's modulus, ρ = density, A = cross-sectional area, I = moment of inertia, w = deflection, P = axial exciting force amplitude, 2ω = exciting frequency, x = axis in the axial direction, and t = time.

Using the reasonable assumption

$$w(x, t)/r = f(t)\sin(\pi x/L) \qquad (2.2)$$

where $r = \sqrt{I/A}$, $f(t)$ = dimensionless deflection at the middle section, L = height of the column, the problem is simplified to the well-known Mathieu equation

$$\Omega^2\frac{d^2 f}{d\tau^2} + (1 + 2\lambda\cos 2\tau)f = 0 \qquad (2.3)$$

where $\tau = \omega t$, $\Omega = \omega/\omega_L$, $\omega_L^2 = \pi^4\mathrm{E}I/\rho AL^4$ – fundamental frequency of tranverse vibration of the column, $\lambda = 1/2\,(P/P_{\mathrm{cr}})$, and $P_{\mathrm{cr}} = \pi^2\mathrm{E}I/L^2$ – the Euler load.

In this particular case, the incremental method can be formulated according to the periodic solution method [3–5] which is based on the fact that the instability boundary of the system has a periodic solution of period $2\pi/\omega$ or $4\pi/\omega$.

Once an instability boundary point (λ_0, Ω_0) corresponding to a periodic solution $f_0(t)$ has been known, then a neighboring boundary point can be written as

$$\begin{aligned}\lambda &= \lambda_0 + \Delta\lambda \\ \Omega &= \Omega_0 + \Delta\Omega \\ f(t) &= f_0(t) + \Delta f(t)\end{aligned} \qquad (2.4)$$

Inserting these expressions into equation (2.3) and retaining only the linear terms of increments, one will obtain

$$\Omega_0^2\frac{d^2\Delta f}{d\tau^2} + (1 + 2\lambda_0\cos 2\tau)\Delta f = R$$

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August, 1981; final revision April, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-2\Delta\lambda f_0 \cos 2\tau - 2\Delta\Omega\Omega_0 \frac{\partial^2 f_0}{\partial \tau^2} \tag{2.5}$$

where

$$R = -\left[\Omega_0^2 \frac{\partial^2 f_0}{\partial \tau^2} + (1 + 2\lambda_0 \cos 2\tau) f_0\right] \tag{2.6}$$

is reserved as a corrective term and the third term on the right-hand side is the unbalanced force due to frequency increment $\Delta\Omega$.

The function $f_0(t)$ and $\Delta f(t)$ are periodic and they can then be taken as the form

$$\left.\begin{aligned} f_0(t) &= \sum_{k=1,3,5,\ldots} (a_k \sin k\tau + b_k \cos k\tau) \\ \Delta f(t) &= \sum_{k=1,3,5,\ldots} (\Delta a_k \sin k\tau + \Delta b_k \cos k\tau) \end{aligned}\right\} \tag{2.7}$$

for solution with period 2π in terms of τ, and

$$\left.\begin{aligned} f_0(t) &= b_0 + \sum_{k=2,4,6,\ldots} (a_k \sin k\tau + b_k \cos k\tau) \\ \Delta f(t) &= \Delta b_0 + \sum_{k=2,4,6,\ldots} (\Delta a_k \sin k\tau + \Delta b_k \cos k\tau) \end{aligned}\right\} \tag{2.8}$$

for the solution with period 4π in terms of τ.

Substituting (2.7) or (2.8) into (2.5) and using the Galerkin procedure

$$\int_0^{2\pi} \left[\Omega_0^2 \frac{d^2 \Delta f}{d\tau^2} + (1 + 2\lambda_0 \cos 2\tau)\Delta f\right] \delta(\Delta f)\, d\tau$$

$$= \int_0^{2\pi} \left[R - 2\Delta\lambda f_0 \cos 2\tau - 2\Delta\Omega\,\Omega_0 \frac{\partial^2 f_0}{\partial \tau^2}\right] \delta(\Delta f)\, d\tau \tag{2.5a}$$

one can easily obtain a set of linear equations of the form

$$[C]\{\Delta a\} = \{r\} + \Delta\lambda\{p\} + \Delta\Omega\{q\} \tag{2.9}$$

where $[C]$ is a coefficient matrix, and $\{\Delta a\}$ is a vector consisting of Fourier coefficients Δa_k or Δb_k, for example.

$$[C] = \begin{bmatrix} 1 + \lambda_0 - \Omega_0^2 & -\lambda_0 & & \\ -\lambda_0 & 1 - (3\Omega_0)^2 & \cdot & \\ & -\lambda_0 & \cdot\ \cdot & \\ & & \cdot\ \cdot\ \cdot & \\ & & \cdot\ \cdot & \end{bmatrix} \tag{2.10}$$

is the matrix for the Fourier coefficients Δa_k in equation (2.7), and

$$\{\Delta a\}^T = [\Delta a_1, \Delta a_3, \Delta a_5, \ldots\ldots] \tag{2.11}$$

$\{r\}$ is the corrective vector derived from (2.6), and $\{p\}$, $\{q\}$ are vectors obtained from the second and third right-hand side terms, respectively. All matrices and vectors are formed automatically by computer.

It can be seen that in contrast with the usual periodic solution method [3], which leads to eigenvalue problems, the present method need only to treat a series of linear algebraic equations. The advantage of this feature will become more pronounced when nonlinearity is taken into account.

Because only the relative values of the coefficients in equations (2.7) or (2.8) are required, it is possible to prescribe one of them as a unity reference constant with its corresponding of increment in $\{\Delta a\}$ set to zero, for example, $a_1 = 1$ and $\Delta a_1 = 0$.

An incremental computation procedure similar to [5] is used. In the solution method, if $\Delta\lambda$ is specified as an active increment and the others (i.e., $\Delta\Omega$ and $\{\Delta a\}$, except for one of the components that has been set to zero) are solved from the linear equations ((2.9) in the present case), then it will be called the λ-incrementation. Similarly, it is equally possible to have the Ω-incrementation, etc. In practice, the active increment is chosen as the one that varies faster.

Figure 1 shows the boundary curves of the principal region of instability computed by λ-incrementation ($\Delta\lambda = 0.1$) with different number of harmonic terms. It can be seen that three harmonic terms will give results of such high accuracy that the curves plotted in Fig. 1 are almost coincident with those from the data of the stability chart for Mathieu's function even in the large value range of λ. To guarantee the numerical accuracy, Newton-Raphson iteration with an updated corrective

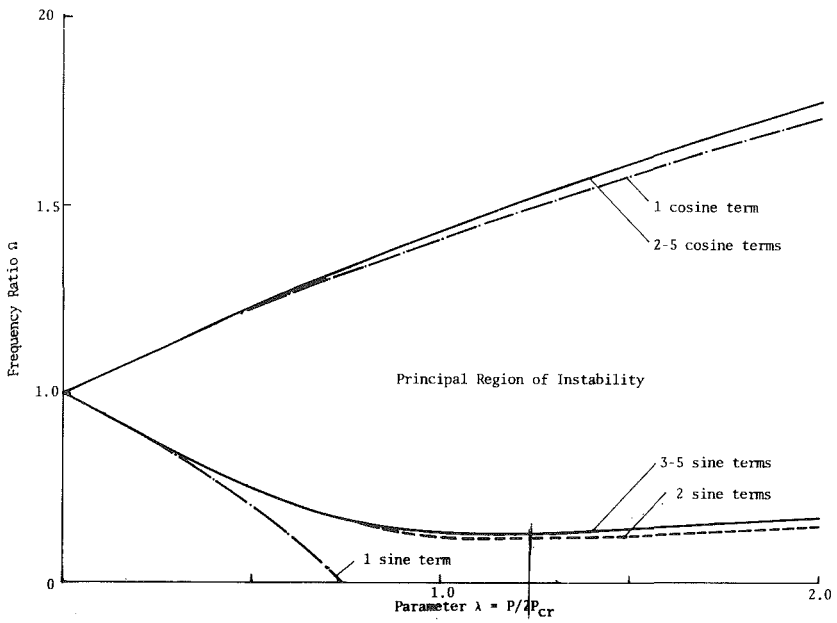


Fig. 1 Principal instability region of hinged-hinged column

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

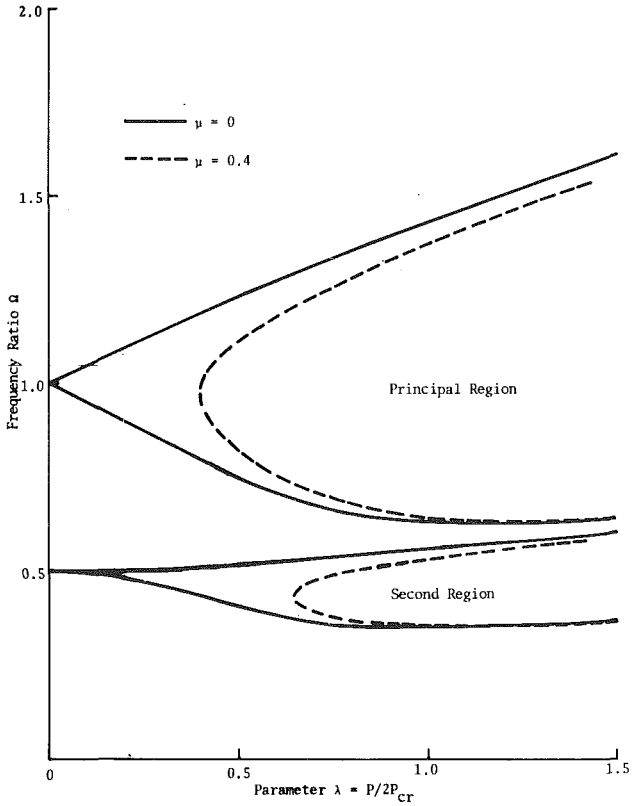


Fig. 2 The influence of damping on principal and second instability regions of hinged-hinged column

vector $\{r\}$ each time within an incremental step has been applied (incremental method with multiple corrections).

This incremental method can be easily generalized to deal with the damped system. If linear viscous damping is assumed, the corresponding differential equation becomes

$$\Omega^2 \frac{d^2 f}{d\tau^2} + \mu\Omega \frac{df}{d\tau} + (1 + 2\lambda \cos 2\tau) f = 0 \tag{2.12}$$

The numerical results for $\mu = 0.4$ are computed and plotted with dotted lines in Fig. 2 (the boundary curves of corresponding undamped system are also plotted for comparison). The computations can be initiated from the solution points of the corresponding undamped system even if the coefficient μ is not small. In fact, a μ-incrementation procedure (with λ fixed) can be used to obtain the necessary damped system solution. In the procedure, the incremental equation (2.5) will be generalized to the following form

$$\Omega_0^2 \frac{d^2 \Delta f}{d\tau^2} + \mu_0 \Omega_0 \frac{d\Delta f}{d\tau} + (1 + 2\lambda_0 \cos 2\tau)\Delta f$$

$$= R - (\Delta\mu\Omega_0 + \mu_0 \Delta\Omega) \frac{df_0}{d\tau} - 2\Delta\lambda f_0 \cos 2\tau - 2\Delta\Omega\Omega_0 \frac{\partial^2 f_0}{\partial\tau^2} \tag{2.13}$$

where

$$R = -\left[\Omega_0^2 \frac{d^2 f_0}{d\tau^2} + \mu_0 \Omega_0 \frac{df_0}{d\tau} + (1 + 2\lambda_0 \cos 2\tau) f_0\right] \tag{2.14}$$

It is noted that the λ-incrementation is not suitable for the U-turn portions of the curves (see Fig. 2) and therefore a capability of automatically changing to Ω-incrementation and vice versa is necessary in the computer program. This case is rather similar to the circumstance near a limit point of nonlinear buckling problem.

It is interesting to point out that the minimum values of the excitation parameters $\lambda_{\min}$, for which the occurrence of undamped vibrations is possible, are predicted according to Bolotin [3] as

$$\lambda_{\min}(k) \cong \sqrt[k]{\mu} \quad (k = 1,2, \ldots)$$

Thus, for the present example, Bolotin gives

$$\lambda_{\min}(1) \cong 0.4$$

for the principal region and

$$\lambda_{\min}(2) \cong 0.6325$$

for the second region. It can be seen from Fig. 2 that $\lambda_{\min}(1)$ is a very good approximation and $\lambda_{\min}(2)$ is also a good approximation but to a lesser degree.

3 Nonlinear Effect on Parametric Instability

A principal advantage of the incremental method is its capability in treating highly nonlinear problems.

Consider again the hinged-hinged column and assume that one of the ends is subjected to a certain axial harmonic motion

$$u_L = \hat{u}_L \cos 2\omega t$$

If the column is undergoing parametric resonance and a large amplitude of deflection is excited, then the problem is described by the following nonlinear equation

$$EI \frac{\partial^4 w}{\partial x^4} - \left(P \cos 2\omega t + \frac{EI}{2L} \int_0^L \left(\frac{\partial w}{\partial x}\right)^2 dx\right) \frac{\partial^2 w}{\partial x^2}$$

$$+ \rho A \frac{\partial^2 w}{\partial t^2} = 0 \tag{3.1}$$

where

$$P = EA\, \hat{u}_L / L. \tag{3.2}$$

By adopting the same mode assumption given in (2.2), equation (3.1) can be reduced to a nonlinear differential equation of the form

$$\Omega^2 \frac{d^2 f}{d\tau^2} + (1 + 2\lambda \cos 2\tau) f + K f^3 = 0 \tag{3.3}$$

where $K = 0.25$ for this particular example.

Based on the procedure in the preceding section, it is possible to arrive at the following incremental equation

$$\Omega_0^2 \frac{d^2 \Delta f}{d\tau^2} + (1 + 2\lambda_0 \cos 2\tau)\Delta f + 3K f_0^2 \Delta f$$

$$= R - 2\Delta\lambda f_0 \cos 2\tau - 2\Delta\Omega\Omega_0 \frac{\partial^2 f_0}{\partial\tau^2}, \tag{3.4}$$

where once again R is a corrective term given by

$$R = -\left[\Omega_0^2 \frac{d^2 f_0}{d\tau^2} + (1 + 2\lambda_0 \cos 2\tau) f_0 + K f_0^3\right] \tag{3.5}$$

For simplicity, only the principal region of instability is considered herein. Let equation (2.7) be substituted into equation (3.4) and then the coefficients of the same harmonic terms are equated as in the harmonic balance method.

A set of linearized equations having the same form as (2.9) is obtained. This can then be solved incrementally by starting from any known linear solution, using amplitude and λ-incrementation procedures as described in the preceding section. Some of the results which are computed by using three harmonic terms to achieve sufficient accuracy, have been plotted in Figs. 3 and 4. It can be seen clearly that the incremental method is not restricted to the case of small parameter and weak nonlinearity.

4 Instability Region of Nonlinear Vibration

One of the important applications of the proposed incremental method is the determination of the instability region of nonlinear vibration.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

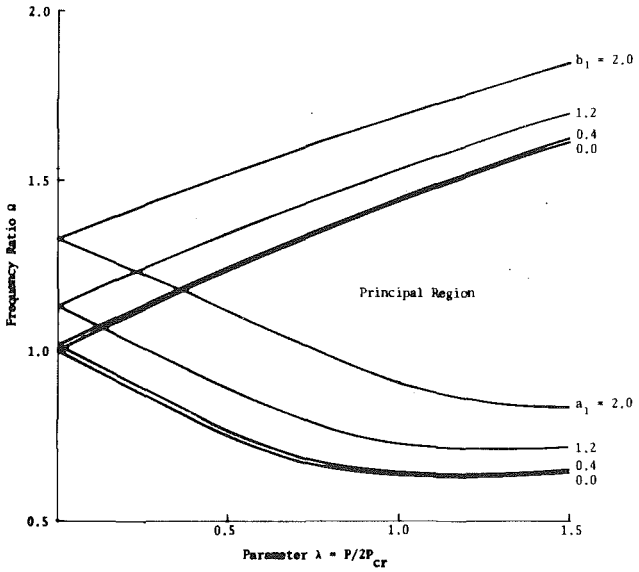


Fig. 3 Nonlinear effect of amplitude on principal instability region

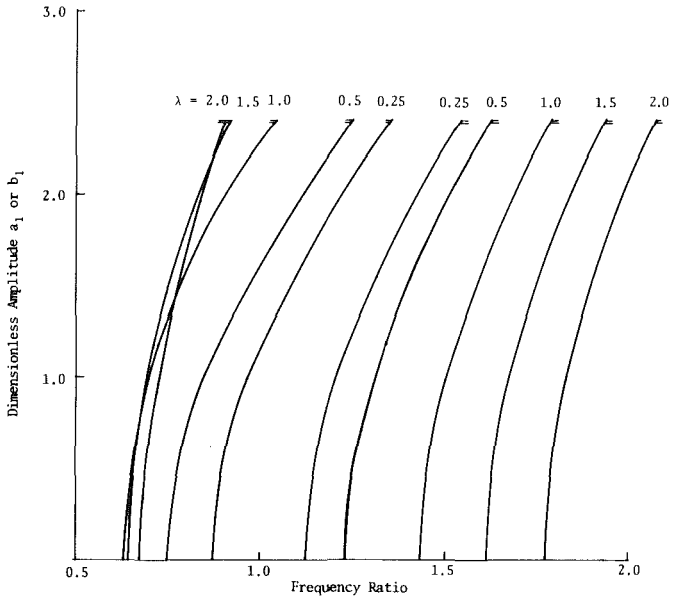


Fig. 4 Nonlinear effect of amplitude on principal instability region. Flags are on the stable sides.

Consider a simply supported beam undergoing superharmonic resonance (of order 3) excited by a central point harmonic load $F\cos\omega t$, This problem has been solved by using the incremental time-space finite strip method in paper [2], where the beam was treated generally as a multiple degree of freedom, but the expression given by (2.2) is still a good approximation for the simply supported beam under resonance condition and its use will simplify the analysis and reduce the governing equation for the present example to (similar to equations (2.3) (3.3))

$$\Omega^2\frac{d^2f}{d\tau^2}+f+Kf^3=F_1\cos\omega t \tag{4.1}$$

where $F_1 = 2FL^3/\pi^4 EIr$.

The instability of the solution for (4.1) can be investigated by considering a small disturbance $\zeta(t)$ in addition to the initial motion $f(t)$, thus if equation (2.2) is modified to

$$w(x,t)/r=(f(t)+\zeta(t))\sin(\pi x/L) \tag{4.2}$$

it follows that equation (4.1) will also be modified to

$$\Omega^2\frac{d^2(f+\zeta)}{d\tau^2}+(f+\zeta)+K(f+\zeta)^3=F_1\cos\omega t \tag{4.3}$$

Recalling that $\zeta(t)$ is small and $f(t)$ satisfies equation (4.1), equation (4.3) can be reduced to

$$\Omega^2\frac{d^2\zeta}{d\tau^2}+\zeta+3Kf^2\zeta=0 \tag{4.4}$$

Since the solution of $f(t)$ is generally expressed as

$$f(t)=\sum_n a_n\cos n\tau \tag{4.5}$$

equation (4.4) is in fact the Hill's equation.

To determine the instability boundary incrementally, assume

$$\left.\begin{aligned} f &= f_0+\Delta f\\ \zeta &= \zeta_0+\Delta\zeta\\ \Omega &= \Omega_0+\Delta\Omega\end{aligned}\right\} \tag{4.6}$$

and substitute them into equation (4.4). When the small quantities of higher order are omitted, one obtains

$$\Omega_0^2\frac{d^2\Delta\zeta}{d\tau^2}+\Delta\zeta+3Kf_0^2\Delta\zeta=R-6Kf_0\zeta_0\Delta f-2\Delta\Omega\Omega_0\frac{d^2\zeta_0}{d\tau^2} \tag{4.7}$$

where the corrective term R has the expression

$$R=-\left[\Omega_0^2\frac{d^2\zeta_0}{d\tau^2}+\zeta_0+3Kf_0^2\zeta_0\right], \tag{4.8}$$

the third right-hand side term is an unbalanced inertia force due to frequency increment $\Delta\Omega$ and the second right-hand side term is an unbalanced force caused by the small change of the state of the initial vibration.

On the instability boundary, $\zeta(t)$ and $\Delta\zeta(t)$ are periodic and assumed to be

$$\zeta(t)=\sum_{n=1,3,5,\ldots}(\alpha_n\sin n\tau+\beta_n\cos n\tau), \tag{4.9a}$$

$$\Delta\zeta(t)=\sum_{n=1,3,5,\ldots}(\Delta\alpha_n\sin n\tau+\Delta\beta_n\cos n\tau), \tag{4.9b}$$

Equation (4.9) together with equation (4.5) are inserted into equation (4.7) and a set of linear equations is obtained by equating the same harmonic terms.

$$[K_T(f_0,\Omega_0)]\{\Delta\gamma\}=\{r(f_0,\zeta_0,\Omega_0)\}+\Delta f\{p(f_0,\zeta_0)\}+\Delta\Omega\{q(\zeta_0,\Omega_0)\} \tag{4.10}$$

where $[K_T]$ = the coefficient matrix, $\{\Delta\gamma\}$ = vector consisted of the coefficients $\Delta\alpha_n$ or $\Delta\beta_n$ $\{r\}$ = corrective vector, $\Delta f\{p\}$ and $\Delta\Omega\{q\}$ are the unbalanced force vectors due to Δf and $\Delta\Omega$, respectively.

Remembering that $f(t)$ is known and that the coefficients a_n can be obtained approximately from [2] (see Fig. 5) and with α_1 or β_1 being fixed to be unity, the a_1-incrementation procedure (with Δa_1 specified as an active increment) can be initiated from a guessed rough-approximate solution. In fact, the corrective vector will make this solution approach the actual one after several iterations.

The computed instability region for the out-of-phase superharmonic resonance of order 3 is shown in the $\Omega-a_3$ plane (not in the $\Omega-a_1$ plane) of Fig. 5.

5 Coupled Multiple Degrees-of-Freedom System

The periodic solution assumption can be applied to the simple parametric resonance of a coupled multiple degrees-of-freedom system. As an example, a hinged-clamped column excited by harmonic axial force $P\cos 2\omega t$ is considered. The lateral deflection is expanded into a series of the form

$$w(x,t)=\sum_i^M f_i(t)\phi_i(x) \tag{5.1}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

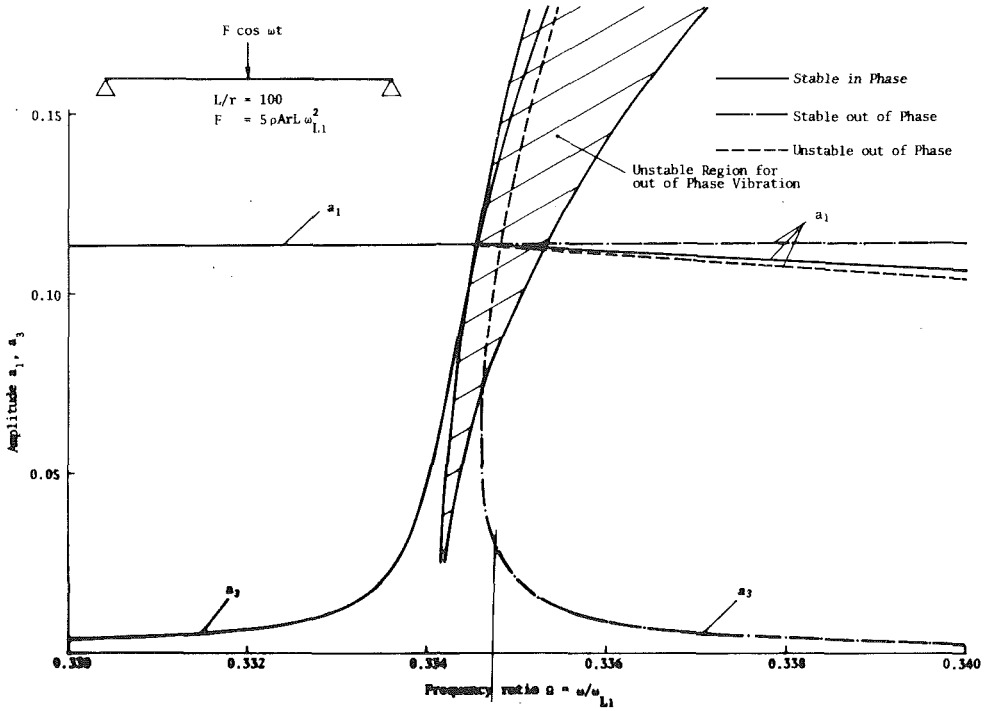


Fig. 5 Instability region of superharmonic resonance of order 3 (out of phase solution). ω_{L1} = linear fundamental frequency, A = cross-sectional area, L = length of beam, r = radius of gyration, and a_1, a_3 = 1st and 3rd harmonic amplitudes of vibration.

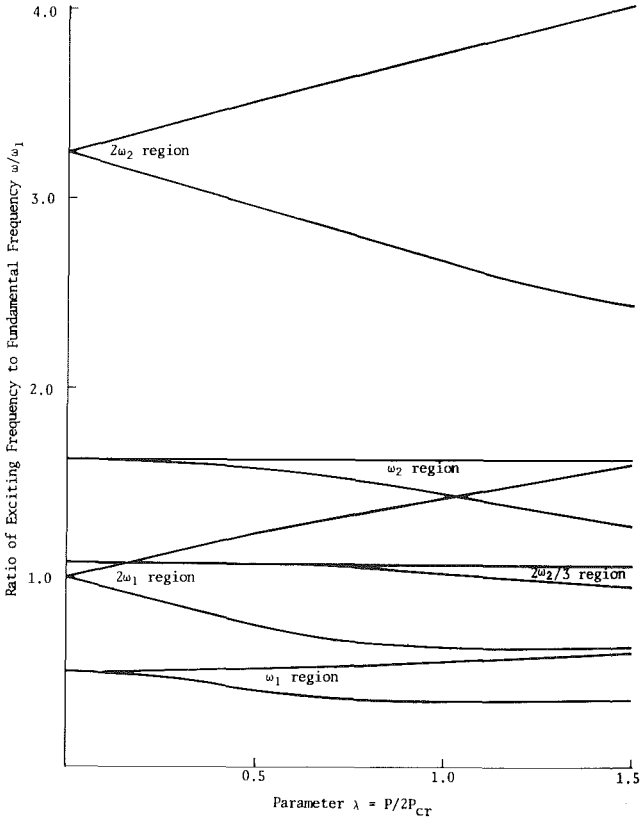


Fig. 6 Instability region of hinged-clamped column

where $\phi_i(x)$ is the ith transverse vibration mode shape of the column and $f_i(t)$ is generally expressed as

$$f_i(t) = \sum_{j=0}^{N} (\alpha_{ij} \sin j\omega t + \beta_{ij} \cos j\omega t) \quad (\alpha_{i0} = 0) \tag{5.2}$$

The same procedures presented in Section 2 are used and the computed instability regions are shown in Fig. 6.

It should be noted that the combination parametric resonance cannot be treated by the periodic solution assumption and an extension of the method [7, 8] or a more general method [9] must be considered. However, this topic will not be dealt with in the present paper although it is believed that the concept of the proposed incremental method is also generally applicable.

In conclusion, it is worthy to mention that a wide variety of problems, linear or nonlinear, in which a series of solutions over a certain range of variation of parameters are needed, could often be analyzed numerically by the present variable parameter incrementation method. Moreover, this method in general has no limitation to small change of parameters and is usually convenient for computer implementation.

References

1 Lau, S. L., and Cheung, Y. K., "Amplitude Incremental Principle for Nonlinear Vibration of Elastic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 959-964.

2 Cheung, Y. K., and Lau, S. L., "Incremental Time-Space Finite Strip Method for Nonlinear Structural Vibrations," *Earthquake Eng. and Struct. Dynamics*, Vol. 10, 1981, pp. 239-253.

3 Bolotin, V. V., *The Dynamic Stability of Elastic Systems*, Holden-Day, San Francisco, 1964.

4 Hsu, C. S., "On Parametric Excitation and Snap-Through Stability Problems of Shells," In: *Thin-Shell Structures*, Fung, Y. C., and Sechler, E. E., Eds., Prentice-Hall, Englewood Cliffs, N.J., 1974.

5 Hu, H. C., "On a Variational Principle in the Theory of Dynamic Stability of Elastic Systems," *Scientia Sinica*, Vol. 23, 1980, pp. 1110-1115.

6 Haisler, W. E., and Stricklin, J. A., "Displacement Incrementation in Nonlinear Structural Analysis by the Self-Correcting Method," *Int. J. Num. Meth. Eng.*, Vol. 11, 1977, pp. 3-10.

7 Lindh, K. G., and Linkins, P. W., "Infinite Determinant Methods for Stability Analysis of Periodic-Coefficient Differential Equations," *AIAA J.*, Vol. 8, 1970, pp. 680-686.

8 Fu, F. C. U., and Nemat-Nasser, S., "Stability of Solution on Systems of Linear Differential Equations With Harmonic Coefficients," *AIAA J.*, Vol. 10, 1972, pp. 30-36.

9 Friedmann, P., and Hammond, C. E., "Efficient Numerical Treatment of Periodic Systems With Application to Stability Problems," *Int. J. Num. Meth. Eng.*, Vol. 11, 1977, pp. 1117-1136.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

T. Irie
Professor.

G. Yamada
Associate Professor.

H. Okada
Student.

Department of Mechanical Engineering,
Hokkaido University,
Kita-13, Nishi-8, Kita-ku,
Sapporo, 060 Japan

Free Out-of-Plane Vibration of a Ring Elastically Supported at Several Points

An analysis is presented for the free out-of-plane vibration of a circular ring elastically supported against deflection, rotation, and torsion at several points located at equal angular intervals. The equations of out-of-plane vibration of the ring is expressed as a matrix differential equation by using the transfer matrix, the solution to which is conveniently given by infinite series. The vibrations arising in the ring are classified into several types, for each of which the natural frequencies and the mode shapes are calculated numerically up to higher modes.

Introduction

This paper presents an analysis of the free out-of-plane vibration of a uniform ring elastically supported against deflection, rotation, and torsion at several points, in which the transfer matrix technique is used. The vibration theory of beams and rings has great importance in many engineering applications such as in the design of machines and structures. The fundamental equations of vibration of curved beams and rings have been presented together with the solution to them in the books of Lord Rayleigh [1], Love [2], and Timoshenko [3]. Formerly, Peterson [4] and Kaiser [5] obtained the natural frequencies of rings experimentally. Recently, Rodriguez [6], Krahula [7], and Rao [8] have treated rings restrained at the whole range, and Culver and Oestel [9], Chen [10], and Petyt and Fleischer [11] have studied multisupported curved beams. However, no papers have been presented for various types of vibration arising in a ring elastically supported against deflection, rotation, and torsion at many points reported here, except for the paper of Rao [12] which treated a ring only transversely supported at several points.

For the purpose of this study, the equations of out-of-plane vibration of a ring are written as a coupled set of first-order differential equations. By introducing the transfer matrix of the ring, these equations are expressed in a matrix differential equation, the solution to which is conveniently written as infinite series. The vibrations arising in rings supported at several points located at equal angular intervals are classified into several types, for each of which the frequency equation is derived by the boundary conditions and the structural symmetry of the ring-spring systems.

By the application of the method, the natural frequencies and the mode shapes of rings are calculated numerically up to higher modes, and the effects of the number and the stiffness of supports and the slenderness ratio of rings on them are discussed.

Analysis

Figure 1 shows a uniform ring elastically supported at several points. With the angular coordinate measured from a support denoted by θ, the X, Y, and Z-axes are taken in radial, transverse, and tangential directions, respectively, as shown in the figure. Without the elongation and shear deformation of the ring taken into account, the equations of out-of-plane vibration of the ring are written as

$$\frac{dQ_Y^*}{d\theta} + \rho A R \omega^2 v^* = 0 \tag{1}$$

$$\frac{dM_X^*}{d\theta} + T^* - RQ_Y^* - \rho I_X \omega^2 \frac{dv^*}{d\theta} = 0 \tag{2}$$

$$\frac{dT^*}{d\theta} - M_X^* + \rho J_Z R \omega^2 \psi = 0 \tag{3}$$

where ρ is the mass per unit volume, R is the radius of curvature, A is the cross-sectional area, I_X and J_Z, respectively, are the second and polar moments of area of the ring, and ω is

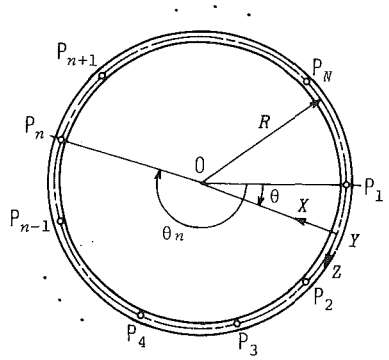


Fig. 1 Uniform ring elastically supported at several points

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June, 1981; final revision, January, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the circular frequency. The bending moment M_X^* and torsional moment T^*, respectively, are given by

$$M_X^* = -\frac{EI_X}{R^2}\left(\frac{d^2v^*}{d\theta^2} - R\psi\right) \tag{4}$$

$$T^* = \frac{GC_Z}{R^2}\left(\frac{dv^*}{d\theta} + R\frac{d\psi}{d\theta}\right) \tag{5}$$

in terms of the transverse deflection v^* and the angle of torsion ψ, where E and G, respectively, are Young's modulus and shear modulus, and C_Z is St. Venant torsional constant of the cross section. The variable Q_Y^* is the shearing force and φ is the angle of rotation expressed as

$$\varphi = -\frac{1}{R}\frac{dv^*}{d\theta} \tag{6}$$

$$\frac{d}{d\theta}\begin{Bmatrix} v \\ \psi \\ \varphi \\ -M_X \\ T \\ Q_Y \end{Bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{\mu} & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & \lambda^4\frac{1}{s_X^2} & 0 & 1 & 0 & -1 \\ 0 & -\lambda^4\left(\frac{1}{s_X^2}+\frac{1}{s_Y^2}\right) & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{Bmatrix} v \\ \psi \\ \varphi \\ -M_X \\ T \\ Q_Y \end{Bmatrix} \tag{11}$$

The continuity conditions and the equilibrium conditions, respectively, at each support ($\theta = \theta_n$) are written as

$$\begin{Bmatrix} v^* \\ \psi \\ \varphi \end{Bmatrix}_{(\theta_n+0)} = \begin{Bmatrix} v^* \\ \psi \\ \varphi \end{Bmatrix}_{(\theta_n-0)},$$

$$\begin{Bmatrix} M_X^* \\ T^* \\ Q_Y^* \end{Bmatrix}_{(\theta_n+0)} + \begin{Bmatrix} M_X^* \\ T^* \\ Q_Y^* \end{Bmatrix}_{(\theta_n-0)} + \begin{Bmatrix} K_R\varphi \\ K_T\psi \\ Kv^* \end{Bmatrix}_{(\theta_n-0)} \tag{7}$$

where K, K_R, and K_T are the translational, rotational, and torsional stiffnesses of a support, respectively, as shown in Fig. 2.

For simplicity of the analysis, the following dimensionless quantities are introduced:

$$v = \frac{v^*}{R}, \quad Q_Y = \frac{R^2}{EI_X}Q_Y^*, \quad (M_X, T) = \frac{R}{EI_X}(M_X^*, T^*) \tag{8}$$

and

$$s_x^2 = \frac{AR^2}{I_x}, \quad s_y^2 = \frac{AR^2}{I_y}, \quad \mu = \frac{GC_z}{EI_x}$$

$$k = \frac{KR^3}{EI_x}, \quad k_R = \frac{K_R R}{EI_x}, \quad k_T = \frac{K_T R}{EI_x} \tag{9}$$

where s_x and s_y are the slenderness ratios of the ring. As a frequency parameter

$$\lambda^4 = \frac{\rho AR^4\omega^2}{EI_x} \tag{10}$$

is used here.

Equations (1)-(6) can be written in a matrix differential equation of the first order

Equation (11) can be written as

$$\frac{d}{d\theta}\{Z(\theta)\} = [M]\{Z(\theta)\} \tag{12}$$

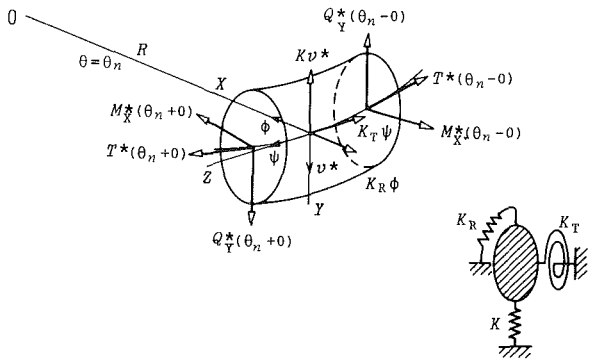


Fig. 2 Force and moment acting at a support

Table 1 Various types of out-of-plane vibration arising in rings elastically supported at equal angular intervals

Number of supports	SS' AA'	AA' AS'	$G(SS')$ $G(AA')$	$G(SA')$ $G(AS')$	SA or AS	$G(SA)$ or $G(AS)$
1	o					
2	o	o				
3	o		o			
4	o	o			o	
5	o		o			
6	o	o	o	o		
7	o		o			
8	o	o			o	o
9	o		o			
10	o	o	o	o		
11	o		o			
12	o	o	o	o	o	o
13	o		o			
14	o	o	o	o		
15	o		o			
16	o	o			o	o
17	o		o			
18	o	o	o	o		

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

by use of the state vector $\{Z(\theta)\} = \{v\ \psi\ \varphi - M_x\ T\ Q_Y\}^T$ and the cross-symmetric coefficient matrix $[M]$ given in (11). In general, $\{Z(\theta)\}$ is expressed as

$$\{Z(\theta)\} = [T(\theta)]\{Z(0)\} \tag{13}$$

by using the transfer matrix $[T(\theta)]$. The substitution of (13) into (12) yields

$$\frac{d}{d\theta}[T(\theta)] = [M][T(\theta)] \tag{14}$$

the solution to which is conveniently expressed as the power series

$$[T(\theta)] = \exp([M]\theta)$$

$$= [I] + \frac{1}{1!}[M]\theta + \frac{1}{2!}[M]^2\theta^2 + \cdots + \frac{1}{n!}[M]^n\theta^n + \cdots \tag{15}$$

Numerical difficulty arises in the calculation of $[T(\theta)]$ given in (15) if the angular interval between adjacent supports is too large. However, it can be overcome by subdividing the interval into several small bays and calculating the transfer matrix for each bay. The entire structure matrix is obtained by assembling the matrices of these bays.

The relations (7) are written as

$$\{Z(\theta_n + 0)\} = [P]\{Z(\theta_n - 0)\} \tag{16}$$

by use of the point matrix

$$[P] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & k_R & 1 & 0 & 0 \\ 0 & k_T & 0 & 0 & 1 & 0 \\ k & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{17}$$

As shown in Table 1, many types of vibration arise in elastic rings supported by many springs of the same stiffness located at equal angular intervals due to the structural symmetry of the ring-spring systems. Here, the symbols S and A, respectively, express symmetrical and antisymmetrical vibrations with respect to a diameter passing through a support, and the symbols S' and A', respectively, express symmetrical and antisymmetrical vibrations with respect to a diameter dividing the angle between adjacent supports equally. Among the variables at the support, we have the relations

$$\{\varphi\ T\ Q_Y\}^T = \{0\ \frac{1}{2}k_T\psi\ \frac{1}{2}kv\}^T \quad \text{for the } S\text{-type vibration} \tag{18}$$

$$\{v\ \psi\ -M_X\}^T = \{0\ 0\ -\frac{1}{2}k_R\varphi\}^T \quad \text{for the } A\text{-type vibration}$$

At the midpoint between adjacent supports, the variables are

$$\{\varphi\ T\ Q_Y\}^T = 0 \quad \text{for the } S'\text{-type vibration}$$

$$\{v\ \psi\ -M_X\}^T = 0 \quad \text{for the } A'\text{-type vibration} \tag{19}$$

In all rings arise the SS' and AA'-type vibrations which are symmetrical and antisymmetrical, respectively, with respect to all diameters passing through the supports and through the midpoints between adjacent supports. The frequency equation of the SS'-type vibration is derived as follows. The state vector at the midpoint $\theta = \pi/N$ between adjacent supports is expressed as

$$\left\{Z\left(\frac{\pi}{N}\right)\right\} = \left[T\left(\frac{\pi}{N}\right)\right]\{Z(+0)\} \tag{20}$$

where N is the number of springs. The substitution of (20) into the first equations of (18) and (19) yields

$$\begin{bmatrix} T_{31} + \frac{1}{2}kT_{36} & T_{32} + \frac{1}{2}k_TT_{35} & T_{34} \\ T_{51} + \frac{1}{2}kT_{56} & T_{52} + \frac{1}{2}k_TT_{55} & T_{54} \\ T_{61} + \frac{1}{2}kT_{66} & T_{62} + \frac{1}{2}k_TT_{65} & T_{64} \end{bmatrix}_{(\pi/N)} \begin{Bmatrix} v \\ \psi \\ -M_X \end{Bmatrix}_{(0)} = 0 \tag{21}$$

with only necessary elements of the transfer matrix $[T(\pi/N)]$ in the calculation. Since the matrix $[M]$ depends on the frequency parameter λ, the transfer matrix $[T(\pi/N)]$ is also a function of λ. Equation (21) gives a set of linear homogeneous equations with unknown coefficients $\{v\ \psi - M_X\}^T_{(0)}$ at a support. For existence of a nontrivial solution of the coefficients, the determinant of the coefficient matrix must be zero. The standard routines cannot be employed for obtaining the eigenvalues of the frequency equation (21). A nontrivial solution is found by searching the values of λ which make the determinant of the coefficient matrix of the equation zero by using Lagrange's interpolation technique. The values of λ, which make the determinant zero, give the eigenvalues of vibration and the mode shapes of the ring determined by calculating the coefficients $\{v\ \psi - M_X\}^T_{(0)}$ corresponding to the eigenvalues.

The frequency equation of the AA'-type vibration is expressed as

$$\begin{bmatrix} T_{13} - \frac{1}{2}k_RT_{14} & T_{15} & T_{16} \\ T_{23} - \frac{1}{2}k_RT_{24} & T_{25} & T_{26} \\ T_{43} - \frac{1}{2}k_RT_{44} & T_{45} & T_{46} \end{bmatrix}_{(\pi/N)} \begin{Bmatrix} \varphi \\ T \\ Q_Y \end{Bmatrix}_{(0)} = 0 \tag{22}$$

which is derived in the aforementioned procedure.

The SA' and AS'-type vibrations arise in rings supported at even-numbered points. The frequency equation of the SA'-type vibration is expressed as

$$\begin{bmatrix} T_{11} + \frac{1}{2}kT_{16} & T_{12} + \frac{1}{2}k_TT_{15} & T_{14} \\ T_{21} + \frac{1}{2}kT_{26} & T_{22} + \frac{1}{2}k_TT_{25} & T_{24} \\ T_{41} + \frac{1}{2}kT_{46} & T_{42} + \frac{1}{2}k_TT_{45} & T_{44} \end{bmatrix}_{(\pi/N)} \begin{Bmatrix} v \\ \psi \\ -M_X \end{Bmatrix}_{(0)} = 0 \tag{23}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and that of the AS'-type vibration is

$$\begin{bmatrix} T_{33}-\frac{1}{2}k_R T_{34} & T_{35} & T_{36} \\ T_{53}-\frac{1}{2}k_R T_{54} & T_{55} & T_{56} \\ T_{63}-\frac{1}{2}k_R T_{64} & T_{65} & T_{66} \end{bmatrix}_{(\pi/N)} \begin{Bmatrix} \varphi \\ T \\ Q_Y \end{Bmatrix}_{(0)} = 0 \quad (24)$$

The $G(SS')$ and $G(AA')$-type vibrations, which are symmetrical and antisymmetrical, respectively, with respect to a diameter passing through at least one of the supports arise in rings supported at odd-numbered points. For these vibrations, we have the following relations in a half part of the rings:

$$\{Z(\pi)\} = \left[T\left(\frac{\pi}{N}\right)\right]\left([P]\left[T\left(\frac{2\pi}{N}\right)\right]\right)^{\frac{N-1}{2}}\{Z(+0)\} = [\bar{T}(\pi)]\{Z(+0)\} \quad (25)$$

$$= \left[\bar{T}\left(\frac{\pi}{2^j}\right)\right]\{Z(+0)\},\ N=2^j(2^i+1)\ (i,j=1,2,\ldots) \quad (26)$$

The frequency equations are also given by replacing the elements of the matrix $[T(\pi/N)]$ of (21)–(24) with those of the structure matrix $[\bar{T}(\pi/2^j)]$ in (26). The eigenvalues of the $G(SS')$, $G(AA')$, $G(SA')$, and $G(AS')$-type vibrations include those of the SS', AA', SA', and AS'-type vibrations, respectively.

The SA and AS-type vibrations arise in rings supported at four, eight, twelve, . . . points, for which we have the following relations between adjacent supports:

$$\left\{Z\left(\frac{2\pi}{N}+0\right)\right\} = [P]\left[T\left(\frac{2\pi}{N}\right)\right]\{Z(+0)\} = \left[\bar{T}\left(\frac{2\pi}{N}\right)\right]\{Z(+0)\} \quad (27)$$

By (18), the frequency equation of the SA-type vibration is expressed as

$$\begin{bmatrix} \bar{T}_{11}+\frac{1}{2}k\bar{T}_{16} & \bar{T}_{12}+\frac{1}{2}k_T\bar{T}_{15} & \bar{T}_{14} \\ \bar{T}_{21}+\frac{1}{2}k\bar{T}_{26} & \bar{T}_{22}+\frac{1}{2}k_T\bar{T}_{25} & \bar{T}_{24} \\ \bar{T}_{41}+\frac{1}{2}k\bar{T}_{46}+\frac{1}{2}k_R\left(\bar{T}_{31}+\frac{1}{2}k\bar{T}_{36}\right) & \bar{T}_{42}+\frac{1}{2}k_T\bar{T}_{45}+\frac{1}{2}k_R\left(\bar{T}_{32}+\frac{1}{2}k_T\bar{T}_{35}\right) & \bar{T}_{44}+\frac{1}{2}k_R\bar{T}_{34} \end{bmatrix}_{(2\pi/N)} \begin{Bmatrix} \upsilon \\ \psi \\ -M_X \end{Bmatrix}_{(0)} = 0 \quad (28)$$

and that of the AS-type vibration is

$$\begin{bmatrix} \bar{T}_{33}-\frac{1}{2}k_R\bar{T}_{34} & \bar{T}_{35} & \bar{T}_{36} \\ \bar{T}_{53}-\frac{1}{2}k_R\bar{T}_{54}-\frac{1}{2}k_T\left(\bar{T}_{23}-\frac{1}{2}k_R\bar{T}_{24}\right) & \bar{T}_{55}-\frac{1}{2}k_T\bar{T}_{25} & \bar{T}_{56}-\frac{1}{2}k_T\bar{T}_{26} \\ \bar{T}_{63}-\frac{1}{2}k_R\bar{T}_{64}-\frac{1}{2}k\left(\bar{T}_{13}-\frac{1}{2}k_R\bar{T}_{14}\right) & \bar{T}_{65}-\frac{1}{2}k\bar{T}_{15} & \bar{T}_{66}-\frac{1}{2}k\bar{T}_{16} \end{bmatrix}_{(2\pi/N)} \begin{Bmatrix} \varphi \\ T \\ Q_Y \end{Bmatrix}_{(0)} = 0 \quad (29)$$

The frequency equations are given only by replacing the elements of the matrix $[T(\pi/N)]$ of (21) and (22) with those of the structure matrix $[\bar{T}(\pi)]$ in (25).

The $G(SS')$, $G(AA')$, $G(SA')$, and $G(AS')$-type vibrations arise in rings supported at even-numbered points, for which we have

$$\left\{Z\left(\frac{\pi}{2^j}\right)\right\} = \left[T\left(\frac{\pi}{N}\right)\right]\left([P]\left[T\left(\frac{2\pi}{N}\right)\right]\right)^i\{Z(+0)\}$$

The $G(SA)$ and $G(AS)$-type vibrations arise in rings supported at the points exceeding four in number, for which we have

$$\left\{Z\left(\frac{\pi}{2^j}+0\right)\right\} = \left([P]\left[T\left(\frac{2\pi}{N}\right)\right]\right)^i\{Z(+0)\} = \left[\bar{T}\left(\frac{\pi}{2^j}\right)\right]\{Z(+0)\},\ N=2^{j+1}i \quad \begin{pmatrix} i=2,3,\ldots; \\ j=1,2,\ldots \end{pmatrix} \quad (30)$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

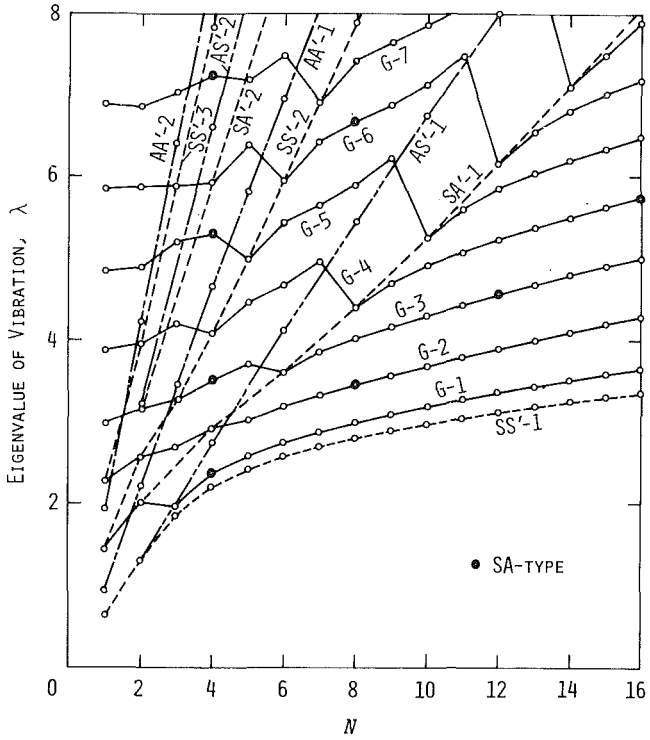


Fig. 3 Eigenvalues of vibration of elastically supported rings versus the number of supports: $\nu = 0.3, s_X = s_Y = 50, k = k_R = k_T = 50$

The eigenvalues of the $G(SA)$ and $G(AS)$-type vibrations also include those of the SA and AS-type vibrations.

Numerical Calculation and Discussion

In this section, the method is applied to rings of uniform circular cross section, which are supported by several springs of the same stiffness located at equal angular intervals, and the eigenvalues of vibration and the mode shapes are calculated numerically up to higher modes.

Figure 3 shows the eigenvalues of vibration of rings versus the number of supports. The eigenvalues of the G-type vibrations not belonging to the SS'-1 vibration and the AA' and AS'-type vibrations are conveniently connected by solid lines. The integer number p attached to the symbol $G-p$ expresses a half of the number of nodes appearing on the rings. The points marked with a double circle ⓞ on the eigencurves of the G-type vibrations express the eigenvalues of the SA and AS-type vibrations, which appear only on the rings supported at four, eight, and twelve points. With an increase in number of supports, the eigenvalues of the G-type vibrations become larger monotonically, except for those of the G-type vibrations exceeding the values of the SA'-1 vibration. Except for the SS'-1 vibration, the eigenvalues of the SS', AA', SA', and AS'-type vibrations become larger rapidly with an increase in number of supports.

Figure 4 shows the eigenvalues of vibration and the mode shapes of a ring supported at six points. The circled points at the left-most figure on the upper row represent the location of

Table 2 Comparison of eigenvalues of vibration λ of free and supported rings with square cross section: $\mu = 1, s_X = s_Y = 4\sqrt{3}$

Circumferential wave number	Free ring	Transversely supported at two points		Clamped at two points	
	Present and Rao [12]	Present	Rao [12]	Present	Rao [12]
2	1.567	1.567	2.124	2.187	2.186
3	2.555	2.555	3.199	3.061	3.059
4	3.416	3.416	4.064	3.854	3.852

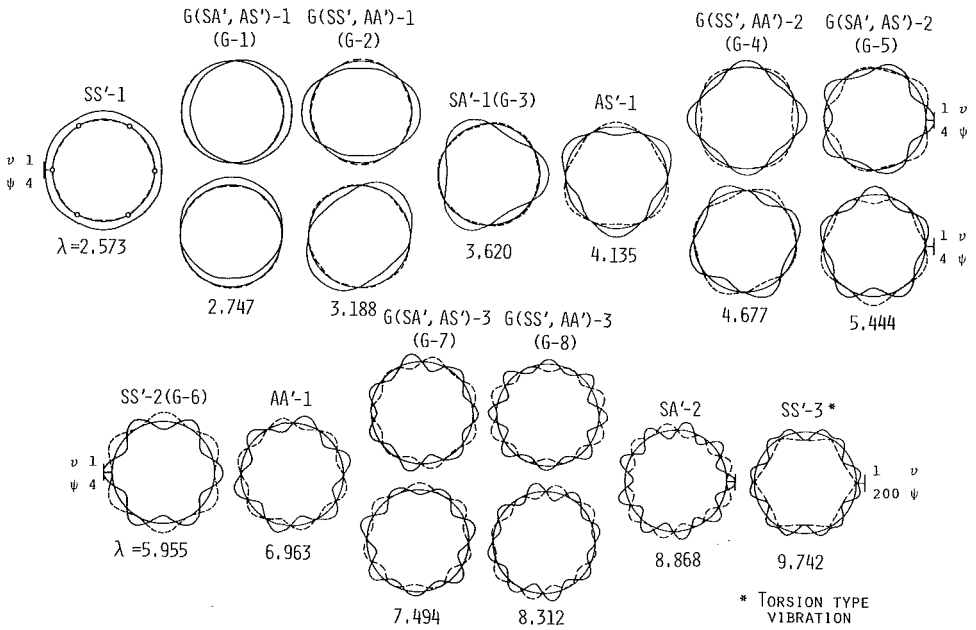


Fig. 4 Mode shapes of a ring elastically supported at six points: $\nu = 0.3, s_X = s_Y = 50, k = k_R = k_T = 50$; ——; ν, – – –; ψ

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

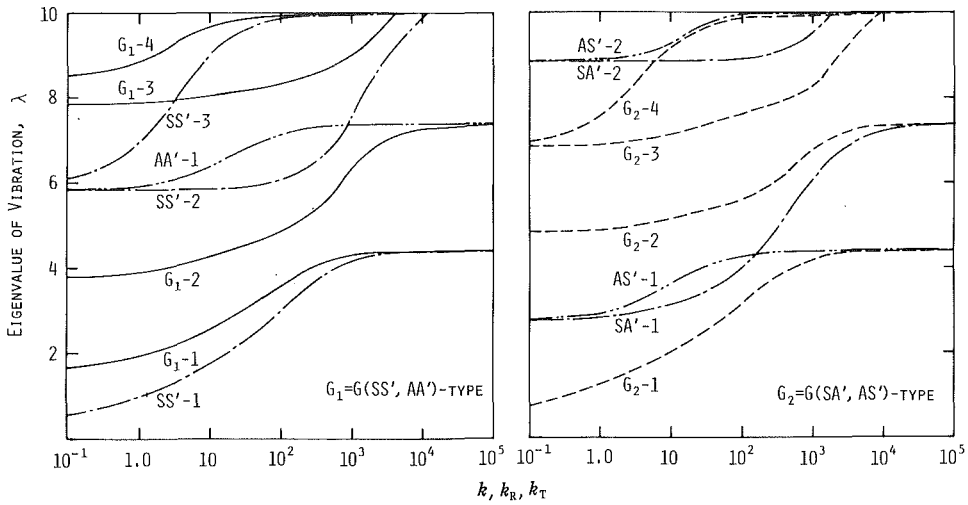


Fig. 5 Eigenvalues of vibration of elastically supported rings versus the stiffness of supports: $\nu = 0.3, s_X = s_Y = 50, N = 6$

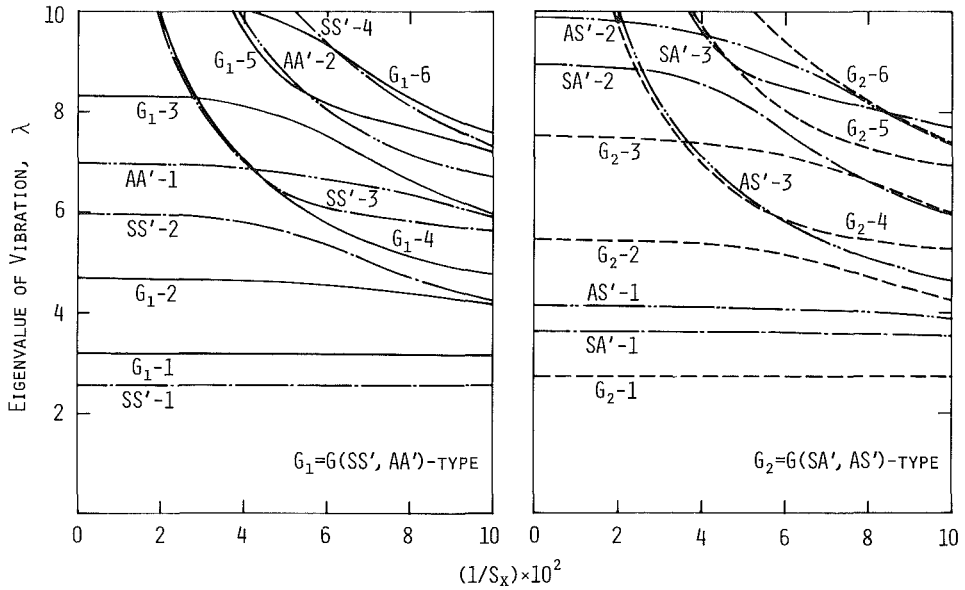


Fig. 6 Eigenvalues of vibration of elastically supported rings versus the slenderness ratio: $\nu = 0.3, N = 6, k = k_R = k_T = 50$

the supports. The set of the $G(SS')$ and $G(AA')$-type vibrations and that of the $G(SA')$ and $G(AS')$-type vibrations have two different mode shapes at the same frequency, which indicates the degeneration of vibration. The SS'-3 vibration marked with * is the torsion-type vibration, in which the angle of torsion is much larger than the transverse deflection in value.

Figure 5 shows the eigenvalues of vibration of rings supported at six points versus the stiffness of supports, where the values of the stiffness are taken to be $k = k_R = k_T$. The ring of $k = k_R = k_T = 0$ is a free ring without supports and that of $k = k_R = k_T = \infty$ is a ring clamped at all supports. With an increase of the stiffness of supports, the eigenvalues of vibration become larger montonically from those of a free ring to those of a clamped ring, although the increasing rate differs depending on the type of vibration. The eigenvalues of the SS'-1 and G_2-1 vibrations become zero for a free ring, which represent the translation and rotation of the ring, respectively. The set of the SS'-2 and AA'-1 vibrations and that of the SA'-1 and AS'-1 vibrations, . . . have the respective same value for a free ring, and the set of the SS'-1 and G_1-1 vibrations and that of the G_2-1 and AS'-1 vibrations, . . . also have the same value for a clamped ring. Frequency crossing arises between the vibrations of different type.

Figure 6 shows the eigenvalues of vibration of rings supported at six points versus the ratio $1/s_X(=1/s_X)$. With an increase of the ratio and hence a decrease of the slenderness ratio s_X, the values become smaller monotonically. Althogh the decreasing rate is different depending on the types of vibration, the rate of the torsion type vibration is larger than that of other type vibrations for the rings of large slenderness ratio.

To illustrate the accuracy of the numerical results obtained here, the eigenvalues of vibration obtained by the present method are compared with the results obtained by Rao taking the rotatory inertia into account [12] for a free ring and a ring transversely supported or clamped at two equally spaced points in Table 2. As shown in the table, the values of a free ring and a clamped ring obtained by the present authors are in good agreement with those of Rao independently of the difference of analytical approach. Although it is natural that the eigenvalues of a transversely supported ring are the same with those of a free ring and are smaller than those of a clamped ring, the values of Rao are different from those of a free ring and the values for $n = 3$ and 4 are higher than those

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of a clamped ring. Therefore, the results of a transversely supported ring by Rao are considered to be incorrect probably due to some mistakes in analytical or computational process.

The numerical computations presented here were carried out on a HITAC M-200H computer of the Hokkaido University Computing Center.

References

1 Rayleigh, J. W. S., *The Theory of Sound,* 2nd ed., Dover Publications, New York, 1945.

2 Love, A. E. H., *A Treatise on the Mathematical Theory of Elasticity,* 4th ed., Dover Publications, New York, 1944.

3 Timoshenko, S., *Vibration Problems in Engineering,* 3rd ed., D. Van Nostrand Company, New York, 1955.

4 Peterson, R. E., "Natural Frequency of Gears," ASME Transactions, Vol. 52 (APM-52), 1930, pp. 1–11.

5 Kaiser, E. R., "Acoustical Vibrations of Rings," *The Journal of the Acoustical Society of America,* Vol. 25, No. 4, 1953, pp. 617–623.

6 Rodriguez, D. A., "Three-Dimensional Bending of a Ring on an Elastic Foundation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, No. 3, 1961, pp. 461–463.

7 Krahula, J. L., "Out-of-Plane Free Vibrations of a Uniform Circular Ring," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, No. 3, 1966, pp. 708–709.

8 Rao, S. S., "Three-Dimensional Vibrations of a Ring on an Elastic Foundation," *Aeronautical Journal,* Vol. 75, 1971, pp. 417–419.

9 Culver, C. G., and Oestel, D. J., "Natural Frequencies of Multispan Curved Beams," *Journal of Sound and Vibration,* Vol. 10, No. 3, 1969, pp. 380–389.

10 Chen, S., "Coupled Twist-Bending Waves and Natural Frequencies of Multispan Curved Beams," *The Journal of the Acoustical Society of America,* Vol. 54, No. 4, 1973, pp. 1179–1183.

11 Petyt, M., and Fleischer, C. C., "Vibration of Multi-Supported Curved Beams," *Journal of Sound and Vibration,* Vol. 32, No. 3, 1974, pp. 359–365.

12 Rao, S. S., "Effects of Transverse Shear and Rotatory Inertia on the Coupled Twist-Bending Vibrations of Circular Rings," *Journal of Sound and Vibration,* Vol. 16, No. 4, 1971, pp. 551–566.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Implicit Numerical Integration for Periodic Solutions of Autonomous Nonlinear Systems

G. A. Thurston
Aerospace Engineer,
NASA Langley Research Center,
Hampton, Va. 23655

A change of variables that stabilizes numerical computations for periodic solutions of autonomous systems is derived. Computation of the period is decoupled from the rest of the problem for conservative systems of any order and for any second-order system. Numerical results are included for a second-order conservative system under a suddenly applied constant load. Near the critical load for the system, a small increment in load amplitude results in a large increase in amplitude of the response.

Introduction

Implicit integration methods are widely used to compute numerical solutions to nonlinear structural dynamics problems. The physical problem is modeled by a set of nonlinear ordinary differential equations. The dependent variables in the differential equations are generalized coordinates or discrete variables. The ordinary differential equations are approximated by difference equations in time, the independent variable, and the nonlinear difference equations are solved iteratively by implicit procedures. For certain problems, however, implicit numerical procedures can be unstable [1], and periodic solutions of nonlinear autonomous systems are inherently difficult to compute accurately by implicit procedures. The purpose of this paper is to present a stable numerical method for computing periodic solutions of nonlinear autonomous systems.

The stable procedure is derived by first recognizing that the linearized difference equations solved at each iteration step by the implicit methods are approximations to the linear variational equations associated with the nonlinear ordinary differential equations. The homogeneous linear variational equations play an important role in the stability theory of nonlinear dynamical systems. The form of the solutions of the linear variational equations for autonomous system is known from Floquet theory. These solutions suggest a change of variables in the differential equations before starting the numerical integration procedure. This paper is mainly concerned with the change of variables independent of the details of the resulting difference equations and their numerical solutions.

Since time t, the independent variable, does not appear explicitly in autonomous systems, the period of vibration is an unknown parameter that depends on the given parameters of the systems and/or the initial conditions. This paper shows how to decouple the computation of the period from the rest of the computations in iterative solutions. The decoupling uses both a change in the independent variable and a change in dependent variables. A change in the independent variable, which makes the product of angular frequency and time a new independent variable, brings the angular frequency into the problem as a parameter to be determined. The change in dependent variables decouples the computation for the angular frequency and also suppresses spurious phase shifts in time.

The change in dependent variables is suggested by a theoretical result, [2, 3]. The result states that in an rth order system of homogeneous linear differential equations if $k \leq r$ linearly independent solutions of the homogeneous equations are known, the solution of the nonhomogeneous rth order system can be obtained by solving $(r-k)$ coupled differential equations plus k equations that can be solved by quadrature.

For autonomous systems with periodic solutions, the decoupled linear variational equations contain a weakly unstable solution which is the product of time and velocity. This unstable solution is removed from the iteration by a proper choice of the angular frequency, [4].

The change in variables is first examined for autonomous second-order equations where analytical results for the linear variational equations can be written concisely. The analysis also serves to sharpen qualitative results from general stability theory. Results using implicit numerical integration are presented for a second-order conservative system under a step load of constant magnitude. New analytical results are also derived for conservative systems with multiple degrees of freedom. These results show how the implicit numerical integration procedure for the second-order system is extended to higher-order systems.

Second-Order Autonomous System

Linear Variational Equation. Consider a second-order equation of motion in the form

$$\ddot{x}+F(x,\dot{x})=0 \tag{1}$$

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February, 1982; final revision, May, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where dots denote differentiation with respect to time t, the independent variable. Since t does not appear explicitly in $F(x,\dot{x})$, the equation is called autonomous.

Implicit numerical integration methods for solving equation (1) discretize the differential equation and solve the resulting nonlinear algebraic equations by an iterative method which is often some version of the Newton-Raphson method. For problems where the solutions of the nonlinear problems are periodic in time, implicit numerical procedures can be unstable, [1].

An alternate procedure is to linearize the nonlinear equation of motion before seeking approximate numerical solutions. The iterative procedure of solving a sequence of linear problems has various names in the literature. The procedure is called Newton's method in this paper. Newton's method applied to equation (1) solves recursively the linear variational equation

$$\delta\ddot{x}_m + p_{m-1}\delta\dot{x}_m + q_{m-1}\delta x_m = -E_{m-1} \tag{2a}$$

plus

$$x_m = x_{m-1} + \delta x_m \quad m = 1,2, \ldots . \tag{2b}$$

where x_{m-1} is a known approximation after $(m-1)$ iterations and

$$p_{m-1} = F_{,\dot{x}}(x_{m-1},\dot{x}_{m-1})$$
$$q_{m-1} = F_{,x}(x_{m-1},\dot{x}_{m-1})$$
$$E_{m-1} = \ddot{x}_{m-1} + F(x_{m-1},\dot{x}_{m-1})$$

A subscript on F following a comma denotes partial differentiation with respect to the subscripted variable. The linear variational equation, (2a), is derived by substituting equation (2b) into equation (1), expanding in a Taylor's series about the current approximate solution x_{m-1}, and retaining linear terms in the correction δx_m.

One method of solving the linear variational equation is to discretize it temporally and solve it numerically. Since Newton's method is a "corrector" method, an implicit numerical integration method is a natural choice for a numerical solution of the linear variational equation. In fact, if the nonlinear equation (1), is discretized by an implicit method and then the resulting discrete equations are linearized by the Newton-Raphson method, the final linear algebraic equations are the same as those generated by an implicit numerical solution of equation (2a). In other words, the operations of linearization and discretization commute.

It is not necessary, however, to solve equation (2a) numerically as it stands. For computing periodic solutions of the nonlinear problem, two changes of variables are proposed here prior to the numerical integration. One is a change in the independent variable and the second is a change in dependent variables. The form of the change in dependent variables is suggested by the limiting case when Newton's method has converged. The limiting case is examined next to show the motivation for the change in variables before presenting numerical results.

Exact Variational Equation Solution. When the Newton iteration of equations (2) converges so that x is the exact solution of the nonlinear equation, the linear variational equation for x is homogeneous and is

$$\delta\ddot{x} + p(x,\dot{x})\delta\dot{x} + q(x,\dot{x})\delta x = 0 \tag{3}$$

Also, when x solves equation (1), one solution of equation (3) is $\delta x = \dot{x}$. This follows directly from differentiating equation (1) with respect to time. Knowing one solution of equation (3), the second-order equation can be decoupled by a change in variables, [2, 3], similar to the method of variation of parameters.

A complete solution of equation (3) is sought in the form

$$\delta x = Ax + B\dot{x} \tag{4a}$$

$$\delta\dot{x} = A\dot{x} + B\ddot{x} \tag{4b}$$

with the condition

$$\dot{A}x + \dot{B}\dot{x} = 0 \tag{5a}$$

In the new variables, $A = A(t)$ and $B = B(t)$, equation (3) becomes

$$\dot{A}\dot{x} + \dot{B}\ddot{x} + G_1(t)A + G_2(t)B = 0 \tag{5b}$$

where

$$G_1(t) = \ddot{x} + p\dot{x} + qx \tag{6a}$$

$$G_2(t) = \ddot{x} + p\ddot{x} + q\dot{x} = 0 \tag{6b}$$

The function $G_2(t)$ vanishes when x is an exact solution of the nonlinear problem and $\dot{x}$ satisfies equation (3). The function is introduced here for use in numerical solutions where x is an approximate solution.

Equations (5) solved for $\dot{A}$ and $\dot{B}$ are

$$W\dot{A} - \dot{x}G_1A = 0 \tag{7a}$$

$$W\dot{B} + xG_1A = 0 \tag{7b}$$

where W is the Wronskian determinant

$$W = x\dot{x} - \dot{x}^2 \tag{8}$$

The derivative of the determinant W is

$$\dot{W} = x\ddot{x} - \dot{x}\ddot{x} \tag{9}$$

and it follows from equation (6b) that

$$-\dot{x}G_1 = \dot{W} + pW \tag{10}$$

Therefore, the solution of equation (7a) is

$$A = C_2 W^{-1}\exp(-\textstyle\int^t p\,d\tau) \tag{11a}$$

Then, B can be determined by a direct integration of equation (7b)

$$B = C_1 - C_2\textstyle\int^t xG_1W^{-2}\exp(-\int^\xi p\,d\tau)\,d\xi \tag{11b}$$

Substitution of A and B into equation (4) completes the solution of equation (3) with arbitrary constants C_1 and C_2.

Diliberto [3] derived a different form for δx starting with

$$\delta x = B(t)\dot{x} + C(t)\ddot{x}$$

He then reviewed classical infinitesimal stability theory in terms of the results. An important special case is the stability when x is periodic in time. Application of the change in dependent variables in equations (4) in computing periodic solutions is examined next.

Iteration for Periodic Solutions. The change of dependent variables in equation (4) is adapted here to the iterative procedure in Newton's method, equation (2), for periodic solutions. The change of the independent variable $\theta = \omega t$ is also introduced. Since ω is an unknown parameter, a correction term $\delta\omega$ is introduced in the iteration in Newton's method. With new unknown functions A_m and B_m plus the unknown constant $\delta\omega_m$, the transformed equations for the mth iteration are

$$W(\theta)A_m' - x'G_1(\theta)A_m - x'G_2(\theta)B_m = x'[E_1(\theta) + \delta\omega_m E_2(\theta)] \tag{12a}$$

$$W(\theta)B_m' + xG_1(\theta)A_m + xG_2(\theta)B_m = -x[E_1(\theta) + \delta\omega_m E_2(\theta)] \tag{12b}$$

$$x_m = x_{m-1} + A_m x_{m-1} + B_m x_{m-1}' \tag{12c}$$

$$\omega_m = \omega_{m-1} + \delta\omega_m \tag{12d}$$

$$E_1 = \omega_{m-1}^2 x_{m-1}'' + F(x_{m-1}, \omega_{m-1}x_{m-1}')$$
$$E_2 = 2\omega_{m-1}x_{m-1}'' + p_{m-1}(x_{m-1}, \omega_{m-1}x_{m-1}')x_{m-1}'$$

Primes denote differentiation with respect to θ. The $(m-1)$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

subscript is implied in the coefficients of A_m and B_m in the first two equations of equations (12), as in

$$W(\theta) = \omega_{m-1}^2 [xx'' - (x')^2]_{m-1}$$

The correction $\delta\omega_m$ on the angular frequency is available to make the current approximation for x periodic. Reference [4] shows that the correction on ω is determined by suppressing a term proportional to $t\dot{x}$ in the solution of the linearized equations. Computing the correction using numerical solutions is now outlined for several cases and numerical results presented for a conservative system under a step loading of constant magnitude.

Conservative System. The simplest example for correcting ω is a conservative system where $F(x,\dot{x}) = F(x)$ and $p = F_{,\dot{x}} = 0$. Assume a periodic solution $x = \phi$ is known for the nonlinear problem that satisfies initial conditions

$$\phi(0) = C \quad \text{and} \quad \phi'(0) = 0$$

Without loss of generality, it can be further assumed that

$$\omega_c = 2\pi/T,\ \phi(\theta + 2\pi) = \phi(\theta) \quad \text{and} \quad \phi(-\theta) = \phi(\theta).$$

The function ϕ is an even function of θ.

To compute a solution with different initial conditions

$$x(0) = C + \Delta C, x'(0) = 0$$

the iteration for Newton's method starts by solving equations (12) with $x_0 = \phi$ and $\omega_0 = \omega_c$ so that $G_2(\theta)$ and $E_1(\theta)$ are zero for the first iteration, $m = 1$. The solution is

$$A_1 = C_2/W + \omega_c \delta\omega_1 (\phi')^2/W \tag{13a}$$

$$B_1 = C_1 + C_2 \int_0^\theta H_1(\xi)\, d\xi - \frac{(2\omega_c \delta\omega_1)}{\omega_c^2} \left(\frac{\theta}{2} + \frac{\phi\phi'}{2(\phi\phi'' - \phi'\phi')} \right) \tag{13b}$$

where

$$H_1(\theta) = -\phi G_1(\theta)/W^2$$

and then,

$$x_1 = \phi + A_1\phi + B_1\phi' \tag{13c}$$

Since the function p is absent, the correction is on ω^2 so that

$$\omega_1^2 = \omega_c^2 + (2\omega_c \delta\omega_1) \tag{13d}$$

replaces equation (12d) for the conservative case.

The correction on ω^2 is determined by enforcing the condition that x_1 is periodic which is equivalent to making B_1 periodic. The integral $H_1(\theta)$ in equation (13b) is an even periodic function so that it has a cosine Fourier series representation

$$H_1(\theta) = a_0 + \sum_n a_n \cos n\theta$$

The integral of $H_1(\theta)$ contains the term $a_0\theta$ so that B_1 is periodic if ω_1^2 is given by

$$(\omega_1^2 - \omega_c^2)/\omega_c^2 = \frac{\Delta C}{C} \frac{W(0)}{\pi} \int_0^{2\pi} H_1(\theta)\, d\theta \tag{14}$$

where the constants of integration satisfying the new initial conditions are $C_1 = 0$ and $C_2 = (\Delta C/C)\, W(0)$. The iteration is continued with $m = 2,3, \ldots$ in equations (12) until E_1 is arbitrarily small. At each iteration, B_m is made periodic by the correct choice of $(\omega_m^2 - \omega_{m-1}^2)$. The functions A_m are even periodic functions of period 2π as long as the implicit integration scheme preserves the symmetry of the x_m. The constant of integration C_1 is set to zero without loss of generality and the solution $C_1 x'$ is easily suppressed in the numerical integration by enforcing the condition $B_m(0) = 0$. The determinant of the transformation, $W(\theta)$, must not vanish in the numerical solution. It vanishes at singular points $x = \bar{x}$, $x' = 0$ where $\bar{x}$ is an equilibrium position, $F(\bar{x}) = 0$, but the linear variational equation has constant coefficients for this case and can be solved without the change of variables. Also, periodic solutions cannot pass through a singular point since this would violate a uniqueness condition, and W is not likely to vanish in practical problems.

A numerical solution for an example of a second-order conservative system has been programmed to illustrate the procedure for correcting ω^2 in Newton's method. The example problem is related to "dynamic buckling" in which a mass on a nonlinear spring is subject to an instantaneously applied step loading of constant magnitude Q. The nondimensionalized equation of motion is

$$\omega^2 y'' + f(y) = Q \quad t \geq 0 \tag{15a}$$

$$y(0) = 0 \tag{15b}$$

$$y'(0) = 0 \tag{15c}$$

To bring the problem into the form of equation (1), let

$$x = y - d$$

$$F(x) = f(x + d) - Q$$

where

$$f(d) = Q \tag{16}$$

The normalized restoring force for the example problem is

$$f(y) = \alpha \sin(\pi y/\sigma) + \beta \sinh y$$

and the slope of the restoring force curve as a function of displacement is

$$F_{,x} = (\alpha\pi/\sigma) \cos(\pi y/\sigma) + \beta \cosh y$$

Cases of particular interest are those where the constants α, β, and σ are such that, $F_{,x}$ takes on negative values, see Fig. 1. In these cases, the equation of motion, equation (15a), has three constant solutions $y = d_i$, $i = 1,2,3$ that satisfy equation (16) for certain values of Q. One critical value of Q, $Q = Q_{cr}$, is defined by the condition that the solution $y(t)$ of equations (15) satisfies

$$\lim_{t \to \infty} y(t) = d_2$$

For $Q \neq Q_{cr}$, the solutions of equations (15) are periodic in time. The numerical solution of equations (12) for the example problem has been programmed for finding the periodic solutions of equations (15). The trapezoidal rule was used for the implicit numerical integration. One cycle of the response curves $y(t)$ for different values of Q are plotted in Fig. 2. The input constants are $\alpha = \beta = 4$ and $\sigma = 3/4$. For these parameters $Q_{cr} = 4.163$ and d_2 for Q_{cr} is $d_2 = 0.6735$.

Once a convergent solution was obtained for a certain Q, that solution was used as the zeroth approximation x_0 in the Newton iteration of equations (12) for a nearby Q. The value

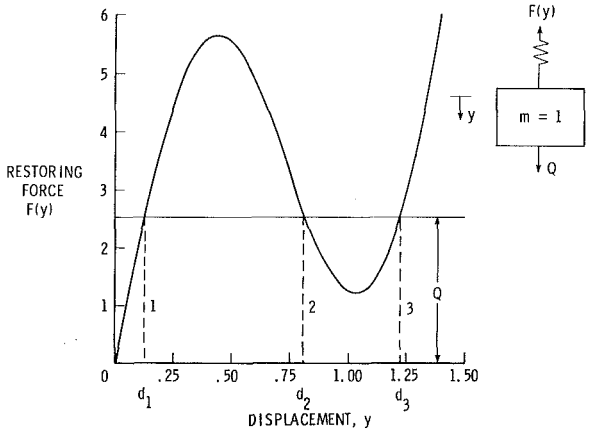


Fig. 1 Nonlinear restoring force $F(y)$ versus y. The displacements d_i are roots of $F(d_i) = Q$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of Q was varied as input data to approach Q_{cr} from above and below. A zeroth first approximation for $Q > Q_{cr}$ is also of the form

$$x_0 = (a - d_2) + a \cos \omega t \tag{17}$$

The method of averaging applied to the residual error leads to the approximation

$$Q = \alpha \sin(\pi a/\sigma) J_0(\pi a/\sigma) + \beta(\sinh a) I_0(a) \tag{18a}$$

$$\omega_0^2 = \frac{2}{a}[\alpha \cos(\pi a/\sigma) J_1(\pi a/\sigma) + \beta(\cosh a) I_1(a)] \tag{18b}$$

where J_0 and J_1 are Bessel functions and I_0 and I_1 are modified Bessel functions. Setting $a = 0.80$ in equations (18) gave $Q = 4.43$ and $\omega_0^2 = 3.83$ for $\alpha = \beta = 4$ and $\sigma = 3/4$. The iteration in equations (12) converged with $\omega^2 = 2.90$ for $Q = 4.43$. Then Q was varied to approach $Q = Q_{cr}$.

As the iteration in equation (12) converged using the trapezoidal rule, the functions $G_1(\theta)$ and $G_2(\theta)$ approached their limiting values. As $x_m \to x = y - d$ the exact solution of equations (15), the function $G_1(\theta) \to -W'/x'$ and $G_2(\theta) \to 0$. It is shown, however, in reference [5] that the rate of convergence of G_1 and G_2 is linear in δx_m while the residual error $E_1(x)$ is quadratic in δx_m. The final accuracy in the converged solutions also depends on truncation and round-off errors. The sample case used numerical harmonic analysis to represent x_m as a cosine series. Then, derivatives of x_m appearing in equations (12) at each integration station were computed from derivatives of the cosine series. Therefore, the residual error and the coefficients in equations (12) have less truncation error than the difference equations for the dependent variables.

The correction on the angular frequency in equation (14) is treated in the numerical solution by superposing two particular solutions of equations (12) for $E_1(\theta)$ and $E_2(\theta)$ so that $B_m(2\pi) = 0$. This correction is a crucial part of the convergence of the iteration for periodic solutions. The importance of the correction is illustrated in Fig. 3. The solid curve is two cycles of the response after one iteration, $m = 1$, of equations (12) for $Q = 4.30$ where x_0 is an exact solution of $Q = 4.43$. The circles are the result of one iteration if time is retained as the independent variable and no numerical correction is made in the period of δx_1. In other words, the circles approximate $x(Q + \Delta Q, \theta)$ by $x(Q) + (\partial x/\partial Q)\Delta Q$. The particular solution $\partial x/\partial Q$ of the linear variational equations is sometimes called the path derivative. The approximation using the path derivative approximates the change in period well over the first half cycle, but then the numerical solution breaks down and starts to oscillate. A term ϵi where ϵ is a constant and i is the station subscript dominates the numerical solution for B_1. The term ϵi is the numerical analog of the term $C_2 a_0 \theta$ that appears in equation (13b) where

$$\pi a_0 = \int_0^{2\pi} H_1(\alpha) d\alpha.$$

The example problem is for a conservative system. The general second-order case has both displacement and velocity-dependent terms. However, periodic solutions can still exist for certain nonlinear problems. Van der Pol's equation [6] is a classic example. Space does not permit discussing such equations in detail except to note that the constants of integration and parameters in the problem can often be adjusted to give periodic solutions in the iteration using Newton's method. The basic idea can also be extended to forced vibrations when a related autonomous system has periodic solutions.

Higher-Order Conservative System

Many methods have been described in the literature for solving nonlinear second-order autonomous systems. Few of the methods are simple to apply to higher-order systems. However, the change in variables of the preceding section can be readily extended to numerical solutions for higher-order conservative systems.

The change in dependent variables for the second-order

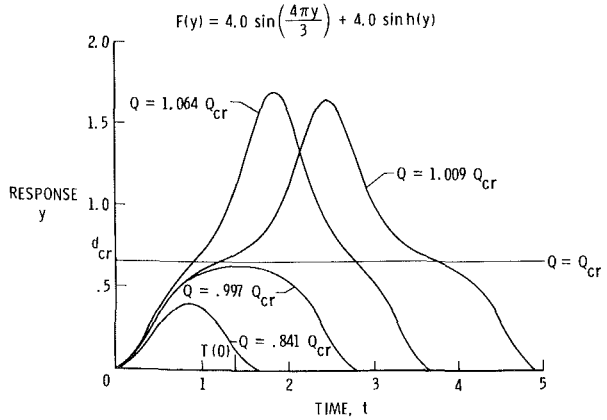


Fig. 2 Response curves for $F(y) = 4.0 \sin(4\pi y/3) + 4.0 \sinh(y)$. $Q_{cr} = 4.163$, $d_{cr} = 0.6735$.

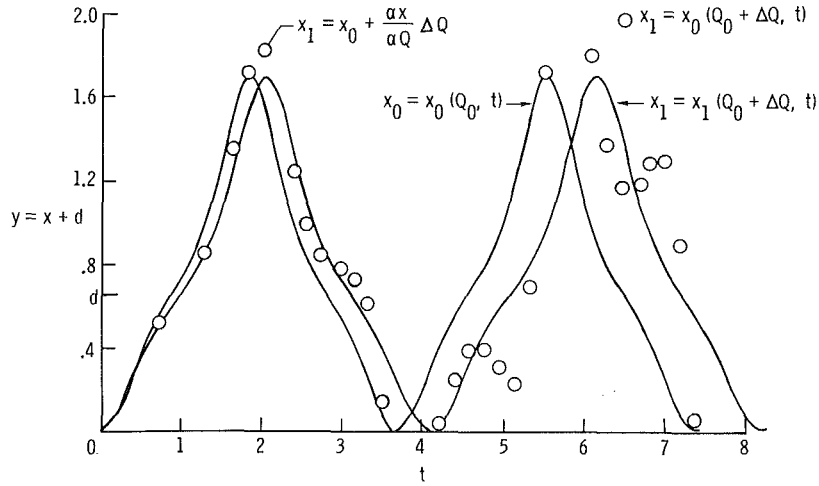


Fig. 3 Results of one interation for solution $x(Q_0 + \Delta Q, t)$ starting with $x_0 = x_0(Q_0, t)$. Solution dx/dQ does not correct period.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

case is based on a solution of the linear variational equation for an exact solution of the nonlinear problem. In this section, it is shown that the linear variational equations for a general conservative system of order $2n$ can be decoupled into a linear system of order $(2n-2)$ and a second system of order 2. In applying Newton's method, the current solution of the nonlinear problem is approximate so that the variational equations are not completely decoupled. Although the coupling terms do not vanish, two solutions can still be identified with a phase shift and with a change in frequency for periodic solutions.

The change in variables for the general case is similar in form to the change in variables for the second-order case except for two new variables. The modification for the new variables is written in subscript notation to show directly the algebraic steps required.

The nonlinear equations of motion for a conservative system of n degrees of freedom are of the form

$$M_i\ddot{x}_i + U_{,i}(x) = 0 \quad i = 1,2,3,\ldots,n \tag{19}$$

The potential energy of the system is $U(x) = U(x_1, x_2, x_3, \ldots, x_n)$ and the subscript i following a comma denotes partial differentiation with respect to x_i. Repeated subscripts are not summed, and dots denote differentiation with respect to time t. After the change in dependent variable $\theta = \omega t$ is introduced, primes denote differentiation with respect to θ.

The iteration in Newton's method for periodic solutions follows along the same line as the case $n = 1$ with

$$x_{im} = x_{i(m-1)} + \delta x_{im} \quad i = 1,2,3,\ldots,n \quad m = 1,2,\ldots \tag{20a}$$

$$\omega_m^2 = \omega_{m-1}^2 + \delta\omega_m^2 \tag{20b}$$

where m is an iteration counter. The subscript m is dropped in subsequent equations. The linear variational equations to determine the δx_i and $\delta\omega^2$ are

$$M_i\omega^2\delta x_i'' + \sum_{j=1}^{n} U_{,ij}\delta x_j = -E_i - \delta\omega^2 M_i x_i'' \quad i = 1,2,3,\ldots,n \tag{21}$$

The terms E_i are the residual error components computed by inserting the current approximation $x_i(\theta)$ and ω^2 into the nonlinear equations, equations (19).

The change of dependent variables for the case $n = 1$ extended to the general case is

$$\delta x_i = A_i x_i + B_i x_i' \tag{22a}$$

$$\delta x_i' = A_i x_i' + B_i x_i'' \tag{22b}$$

with the condition

$$A_i' x_i + B_i' x_i' = 0 \quad i = 1,2,3,\ldots,n \tag{22c}$$

In terms of the new dependent variables, A_i and B_i, the linear variational equations (21), plus equations (22c) are arranged as

$$W_i A_i' - x_i'\left[M_i\omega^2(A_i x_i'' + B_i x_i''') + \sum_{j=1}^{n} U_{,ij}(A_j x_j + B_j x_j')\right] = x_i' E_i + \delta\omega^2 M_i x_i' x_i'' \quad i = 1,2,3,\ldots,n \tag{23a}$$

$$W_i B_i' + x_i\left[M_i\omega^2(A_i x_i'' + B_i x_i''') + \sum_{j=1}^{n} U_{,ij}(A_j x_j + B_j x_j')\right] = -x_i E_i - \delta\omega^2 M_i x_i x_i'' \quad i = 1,2,3,\ldots,n \tag{23b}$$

where

$$W_i = M_i\omega^2[x_i x_i'' - (x_i')^2] \tag{24}$$

Another change of variable is required to identify the correct procedure to fix $\delta\omega^2$ so that the δx_i are periodic. Before writing the transformations in detail, the general approach is outlined by writing the problem in matrix notation. The $2n$ first-order linear equations (23) can be cast into the general form

$$y' = G(\theta)y + E_1(\theta) + \delta\omega E_2(\theta) \tag{25}$$

by dividing the ith equation by W_i. The transpose of the solution vector y is

$$y^T = [A_1, A_2 \ldots A_n, B_1 \ldots B_n]$$

The transformation (see reference [2] on reduction of the order of a homogeneous system)

$$y = Tz \tag{26}$$

carries equations (25) into

$$z' = T^{-1}(GT - T')z + T^{-1}(E_1 + \delta\omega^2 E_2) \tag{27}$$

If k columns of the transformation matrix T are independent solutions of the homogeneous system for equations (25), k components of z do not appear on the right side of equation (27). In this case $(2n-k)$ equations are solved as a coupled system and the remaining k equations by direct integration.

The transformation in equations (4) for $2n = 2$ is an example of equation (31) with $k = 1$ and

$$T = \begin{bmatrix} x & \ddot{x} \\ \dot{x} & \ddot{x} \end{bmatrix} \tag{28}$$

The integration in equation (13b) is an example of a direct integration to determine a particular solution. The shift in the square of the angular frequency $\delta\omega^2 = 2\omega\delta\omega$ is determined by equation (13b) to keep the current approximate solution of the nonlinear problem periodic.

Equations similar to equations (13) are derived here for the general case of order $2n$ with $k = 2$. The practical problem is determining the transformation matrix T and its inverse T^{-1}. In the procedure outlined here, the $k = 2$ columns of the T matrix are approximate but improve as the iteration in Newton's method converges.

The additional change of variables analogous to equation (31) in subscript notation is

$$B_i = b_i + b_I \quad i \neq I \tag{29a}$$

$$B_i = b_I \quad i = I \quad 1 \leq I \leq n \tag{29b}$$

$$A_i = a_i \quad i \neq I \tag{29c}$$

$$A_I = \left[a_I - \sum_{i \neq I}^{n} W_i a_i + \delta\omega^2 \sum_{i=1}^{n} (x')^2\right] / W_I \tag{29d}$$

The matrix T is sparse so that actually only two new variables are introduced, the rest of the variables are merely renamed. The new variables are a_I and b_I where I is arbitrary. The transformation in equations (29) is chosen to decouple equations (23) when the x_i satisfy equations (19). The decoupling can be shown by direct substitution of equation (29) into equation (23) and using two additional relations. The first relation follows from differentiating each of equations (19) with respect to θ,

$$M_i\omega^2 x_i''' + \sum_{i=1}^{n} U_{,ij}x_j' = 0 \quad i = 1,2,3,\ldots n \tag{30}$$

In the new variables, a_i and b_i, the coefficients of b_I vanishes in each equation because of equation (30) while b_I' appears in each equation of the new set corresponding to equations (23b). In the Ith equation, b_I' is the only derivative term so this equation can be solved for b_I' in terms of the other variables and b_I' can then be eliminated from the remaining differential equations.

The second relation of importance, when equations (19) and equations (30) are satisfied, is that the function a_I in equation (29d) is a constant. To show this property of a_I, sum equations (23a) on i, interchange the order of summation on i

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and j, note that $U_{,ij} = U_{,ji}$, use equation (30), and integrate to give

$$\sum_{i=1}^{n} W_i A_i = a_I + (\delta\omega^2/2) \sum_{i=1}^{n} M_i (x_i')^2 \tag{31}$$

where a_I is the constant of integration. The constant of integration a_I depends on initial conditions and the value of $\delta\omega^2$. Equation (35) is the inverse relation for equation (29*d*) and reduces to equation (13*a*) with $a_I = C_2$ for the special case $n = 1$.

After the substitution of equations (29) into equations (23) and the elimination of b_I, the remaining $(2n-2)$ equations are solved for a_i and b_i in terms of a_I and $\delta\omega^2$. Finally, the expression for b_I' in the Ith equation of equation (23*b*) is integrated. The integrated equation corresponds to equation (13*b*).

The integration of b_I' introduces a new constant of integration. This constant is suppressed to eliminate unwanted phase shifts in the δx_i functions. A term containing $\delta\omega^2$ is available to make b_I periodic in θ.

The condition that b_I be periodic is a necessary condition for the convergence of Newton's method to a unique periodic solution of the nonlinear problem. Sufficient conditions for convergence depend on the stability of the reduced system of linear variational equations of order $(2n-2)$, [7, 8].

The changes in variables in equations (22) and equations (29) provide a practical method for isolating the effect of a small change in phase in time and of a change in period from the rest of the computations for periodic solutions of nonlinear autonomous systems.

Concluding Remarks

A change of variables has been derived that stabilizes implicit numerical integration methods for periodic solutions of autonomous systems. Unwanted phase shifts in time are suppressed directly in the numerical solution. A weakly unstable solution of the linear variational equations for the nonlinear system is shown to be related to a change in period. A shift in the period is computed to eliminate the unstable solution. For conservative systems of any order, the change of variables decouples the computation of the period from the rest of the numerical integration.

The results for the special case of a second-order linear variational equation are particularly instructive since the analysis is concise. For any autonomous system of second order, a second solution, in addition to the known solution $\dot{x}$, is derived for the variational equation. The second solution is obtained by quadrature where the integral is a proper integral that is readily evaluated numerically. The integral can also be examined qualitatively to clarify concepts from stability theory. Numerical results are presented in the paper for a second-order conservative system under a suddenly applied constant load.

The frequency computation for autonomous systems is an application of a general computational strategy for Newton's method outlined in reference (5). The procedure is applicable whenever homogeneous solutions satisfying boundary conditions and/or initial conditions exist for the linear variational equations of the problem. Convergence of Newton's method can often be accelerated by fixing amplitudes of these homogeneous solutions and solving for free parameters, such as load parameters or frequency parameters.

Acknowledgment

N. J. Bains programmed the numerical solution for the example case.

References

1 Park, K. C., "An Improved Stiffly Stable Method for Direct Integration of Nonlinear Structural Dynamics Equations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, June 1975, pp. 464–470.

2 Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955, Chapter 3, Section 2.

3 Diliberto, S. P., "On Systems of Ordinary Differential Equations," *Contributions to the Theory of Nonlinear Oscillations*, Annals of Mathematical Studies, No. 20, Princeton University Press, Princeton, N.J., 1950, pp. 1–38.

4 Thurston, G. A., "Floquet Theory and Newton's Method," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, Dec. 1973, pp. 1091–1096.

5 Thurston, G. A., "Newton's Method: A Link Between Continuous and Discrete Solutions of Nonlinear Problems," *Research in Nonlinear Structural and Solid Mechanics*, NASA CP-2147, 1980, pp. 27–46.

6 van der Pol, B., "Forced Oscillations in a System With Nonlinear Resistance," *Philosophical Magazine*, Vol. 3, No. 7, 1927, pp. 65–80.

7 Goodier, J. N., and McIvor, I. K., "The Elastic Cylindrical Shell Under Nearly Uniform Radial Impulse," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, June 1964, pp. 259–266.

8 Hubka, W. F., "Dynamic Buckling of the Elastic Cylindrical Shell Subjected to Impulsive Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 401–406.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

D. J. Inman
Assistant Professor,
Department of Mechanical and
Aerospace Engineering,
State University of New York
at Buffalo,
Buffalo, N.Y. 14260
Assoc. Mem. ASME

A. N. Andry, Jr.
Associate Professor,
Department of Mechanical Engineering,
Michigan State University,
East Lansing, Mich. 48824

The Nature of the Temporal Solutions of Damped Distributed Systems With Classical Normal Modes

Conditions under which the time response of certain distributed parameter systems, which are assumed to possess "classical normal" modes, is critically damped, overdamped, or underdamped are presented. The conditions are derived from the definiteness of certain combinations of the coefficient operators of the describing equations. These conditions are compared to previous results and their usefulness is illustrated by examples.

1 Introduction

One of the most elementary problems in linear vibration theory concerns the motion of a spring-mass-damper system (spring and damper in parallel connecting the mass to a rigid reference point). The qualitative nature of the solution of such systems follows from a short calculation of the critical damping coefficient. A comparison with this value to given values of the damping coefficient determines whether the system is overdamped, critically damped, or underdamped.

Similar calculations are available for lumped parameter multidegree-of-freedom systems. The notions of critical damping, overdamping, and underdamping become more complex in this case but several results are available. One of the first efforts was that of Duffin [1] who defined overdamped systems in terms of a function of quadratic forms of the coefficient matrices of the lumped model. Nicholson [2] defined underdamping in terms of the eigenvalues of these matrices. Müller [3] derived a sufficient condition for underdamping in terms of the definiteness of the coefficient matrices.

Recently, additional results have been obtained by Beskos and Boley [4]. The techniques used in [4] are different from those taken from matrix theory which were used in [1-3]. Essentially, Beskos and Boley address the following problem: given $\{c_k\}_{k=1}^{n}$ different damping coefficients, find the locus of combinations of c_k leading to critically damped motion and thus separating regions of partial or complete overdamping from those of underdamping. Even though the method presented is quite general, its present implementation is limited to small-order systems (due to computational load).

Sharper and more complete results for the lumped parameter case are due to Inman and Andry [5]. The definition of critical damping was stated in terms of the coefficient matrices in a manner analogous to the single degree-of-freedom systems, i.e., a critical damping matrix, C_{cr}, is defined in terms of the unique positive definite square roots of the mass matrix M and the stiffness matrix K to be $C_{cr} = 2(M^{-1/2} K M^{-1/2})^{1/2}$. Having done this, the difference between the critical damping matrix and the actual damping matrix is studied and a system is defined as critically damped, overdamped, or underdamped if the matrix $(C - C_{cr})$ is zero, is positive definite, or is negative definite, respectively. There is a fourth possibility for the multidegree-of-freedom case: $(C - C_{cr})$ is indefinite, and this case is also considered. In each case, the solutions were studied from a qualitative standpoint [5].

Beskos and Boley have also extended their results to the distributed parameter case [6]; their results are again limited by computational load. However, their general approach is not restricted to systems that possess classical normal modes as the work presented here is. In addition they have developed a finite element method based on the concept of the dynamic stiffness coefficients. This method is capable of treating beam structures with different amounts of damping per structural member.

2 Problem Definition

We wish to consider systems governed by the following:

$$u_{tt}(x,t) + L_1[u_t(x,t)] + L_2[u(x,t)] = 0 \quad \text{in} \quad \Omega \qquad (2.1)$$

$$B[u(x,t)] = 0 \quad \text{on} \quad \partial\Omega \qquad (2.2)$$

where

a $(\cdot)_t$ indicates partial differentiation of $(\cdot)$ with respect to t;

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December, 1981; final revision, March, 1982. Paper No. 82-WA/APM-14.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

b Ω is a bounded, open region in R^n, $n = 1, 2$, or 3 with boundary $\partial\Omega$;

c L_1, L_2 are linear spatial differential operators of order n_1, n_2, respectively, and are self-adjoint[1] on the domain $\mathfrak{D}(L) = \{u(\cdot, t) \in L^2(\Omega)$ such that all partial derivatives with respect to x of order up to and including k are in $L^2(\Omega)$ where $k = \max(2n_1, n_2)$ and $B[u(x,t)] = 0$ for $x \in \partial\Omega$, for all $t \geq 0.\}$;

d L_1 and L_2 commute on $\mathfrak{D}(L)$; L_1, L_2 are positive definite operators on $\mathfrak{D}(L)$ and each has a compact resolvent;

e B is a linear operator (and may be a differential operator of order up to max $(n_1, n_2) - 1$) which reflects boundary conditions.

With additional assumptions equations (2.1) and (2.2) can adequately describe vibration problems related to damped beams, plates, shells, etc.

Caughey and O'Kelly [7] have shown that if L_1 and L_2 are self-adjoint and commute on $\mathfrak{D}(L)$ then the system (2.1), (2.2) has a solution given by

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t)\phi_n(x) \tag{2.3}$$

where the $\phi_n(x)$ are a complete set of eigenfunctions (called *classical modes*) of L_1 and L_2. It is tacitly assumed in [7] and here that the functions $a_n(t)$ are at least twice differentiable with respect to t. The set of $\{\phi_i(x)\}$ are elements of $\mathfrak{D}(L)$. Under these assumptions, this paper derives necessary and sufficient conditions for the temporal functions $a_n(t)$ to be periodic (or aperiodic). These conditions are stated in terms of the definiteness of certain combinations of the coefficient operators L_1 and L_2 and are easy to check. Furthermore they yield relationships among the stiffness, damping, and mass parameters of a given system which may be used as design criteria to obtain a desired type of response.

This work is an extension to distributed systems of similar theorems for lumped parameter systems recently derived by the authors [5]. Examples of the utility of the results are given, along with their derivation and a comparison to other results in the current literature.

Following the various possibilities for damping defined for multiple degree-of-freedom systems it is natural to classify systems in the following way: the system described by equations (2.1), (2.2) is called

a *critically damped* if the operator $L_1^2 - 4L_2$ is zero on $\mathfrak{D}(L)$;

b *overdamped* if the operator $L_1^2 - 4L_2$ is positive definite on $\mathfrak{D}(L)$;

c *underdamped* if the operator $4L_2 - L_1^2$ is positive definite on $\mathfrak{D}(L)$;

d *mixed damped* if the operator $L_1^2 - 4L_2$ is indefinite on $\mathfrak{D}(L)$.

3 Main Results

With the preceding classifications and the assumption that the systems of interest possess classical normal modes, it will be shown that qualitative statements can be made about the temporal solutions $a_n(t)$ in (2.3) based on the definiteness of the operator $L_1^2 - 4L_2$ on $\mathfrak{D}(L)$. These statements are direct analogies of the single degree-of-freedom lumped mass case.

Critically Damped Systems. First it will be shown that, if (2.1) is critically damped, then all of the functions $a_n(t)$ in the series solution (2.3) are critically damped functions of time. To illustrate this, substitute (2.3) into (2.1) to yield

$$\sum_{n=1}^{\infty} \left(\ddot{a}_n(t)\phi_n(x) + \dot{a}_n(t)L_1[\phi_n(x)] + a_n(t)L_2[\phi_n(x))\right] = 0. \tag{3.1}$$

Since $L_1^2 = 4L_2$ on $\mathfrak{D}(L)$, this becomes

$$\sum_{n=1}^{\infty} \left(\ddot{a}_n(t)\phi_n(x) + \dot{a}_n(t)L_1[\phi_n(x)] + \frac{a_n(t)}{4}L_1^2[\phi_n(x))\right] = 0. \tag{3.2}$$

Taking the set $\{\phi_i(x)\}_{i=1}^{\infty}$ to be the eigenfunctions of L_1, we obtain

$$L_1\phi_n(x) = \lambda_n\phi_n(x),$$
$$L_1^2\phi_n(x) = \lambda_n^2\phi_n(x),$$

where λ_n represents the eigenvalue associated with $\phi_n(x)$. Taking the obvious inner product between the left-hand side of (3.2) and $\phi_m(x)$, $m \neq n$, and using the orthonormality assumption yields

$$\ddot{a}_n(t) + \lambda_n\dot{a}_n(t) + \frac{\lambda_n^2}{4}a_n(t) = 0 \quad \text{for all} \quad n. \tag{3.3}$$

Solving this for $a_n(t)$ yields

$$a_n(t) = (C_n t + D_n)e^{-\frac{\lambda_n}{2}t}$$

since the discriminant associated with (3.3) is zero. The C_n and D_n are constants determined by the initial conditions. L_1 is assumed positive definite so that $\lambda_n > 0$ for all n. This implies $a_n(t)$ is a critically damped function of time for each n.

Overdamped Systems. If equation (2.1) is overdamped then all of the functions $a_n(t)$ in the series solution (2.3) are overdamped functions of time and the time response does not oscillate.

To see this result note first that since $L_1L_2 = L_2L_1$ on $\mathfrak{D}(L)$, the operators L_1 and L_2 share a common set of eigenfunctions (see [7] for a proof) and that the set is complete since L_1 and L_2 are assumed to have compact resolvents. Hence equation (3.1) in this case is written as

$$\sum_{n=1}^{\infty} \{\ddot{a}_n(t)\phi_n(x) + \dot{a}_n(t)\lambda_n^{(1)}\phi_n(x) + a_n(t)\lambda_n^{(2)}\phi_n(x)\} = 0 \tag{3.4}$$

where $\lambda_n^{(1)}$ is the eigenvalue of L_1 associated with $\phi_n(x)$ and $\lambda_n^{(2)}$ is the corresponding eigenvalue of L_2. Again taking the usual inner product of the left-hand side of (3.4) with $\phi_n(x)$ and using the orthogonality of the eigenfunctions yields

$$\ddot{a}_n(t) + \lambda_n^{(1)}\dot{a}_n(t) + \lambda_n^{(2)}a_n(t) = 0 \quad \text{for all} \quad n. \tag{3.5}$$

Trivially, the characteristic equation of (3.5) is

$$r_n^2 + \lambda_n^{(1)}r_n + \lambda_n^{(2)} = 0$$

The roots of this equation will be negative, real, and unequal if and only if

$$[\lambda_n^{(1)}]^2 - 4\lambda_n^{(2)} > 0 \quad \text{for all} \quad n. \tag{3.6}$$

We claim that if $L_1^2 - 4L_1 > 0$ on $\mathfrak{D}(L)$, then (3.6) holds.

In the foregoing, note that since $L_1^2 - 4L_2$ is assumed to be positive definite on $\mathfrak{D}(L)$, its point spectrum is positive. Since L_1 and L_2 have the same eigenfunctions these eigenfunctions are also the eigenfunctions of the operator $L_1^2 - 4L_2$. Thus the eigenvalues of $L_1^2 - 4L_2$ are found from

$$(L_1^2 - 4L_2)[\phi_n(x)] = L_1^2[\phi_n(x)] - 4L_2[\phi_n(x)]$$
$$= \{[\lambda_n^{(1)}]^2 - 4\lambda_n^{(2)}\}\phi_n(x).$$

[1] n_1, n_2 are necessarily even numbers.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Thus, since $L_1^2 - 4L_2$ has a positive spectrum $[\lambda_n^{(1)}]^2 - 4\lambda_n^{(2)}$ must be positive for all values of n. Hence (3)–(6) yields that each $a_n(t)$ is a nonoscillating decaying function of time of the form

$$a_n(t) = C_n e^{r_n^{(1)}t} + D_n e^{r_n^{(2)}t}$$

where $r_n^{(1)}$, $r_n^{(2)}$ are the roots of the characteristic equation for each n.

Underdamped Systems. If equation (2.1) is underdamped, then all of the functions $a_n(t)$ in the series solution of (2.3) are underdamped functions of time and the time response consists of decaying oscillations.

The proof of this result follows readily from the details of the overdamped case. If $4L_2 - L_1^2$ is positive definite on $\mathfrak{D}(L)$ and if L_1 and L_2 commute, then the eigenvalues of $4L_2 - L_1^2$ become $4\lambda_n^{(2)} - [\lambda_n^{(1)}]^2$, which are all positive. This makes the discriminate (3.6) negative for all n so that each $a_n(t)$ will be sinusoidal with exponentially decaying amplitude. The amplitude decays since L_1 is positive definite so that $-\lambda_n^{(1)}/2$ is negative for all n. Thus, the solutions are of the form

$$a_n(t) = e^{-(\lambda_n/2)t}(C_n \sin \omega_n t + D_n \cos \omega_n t)$$

Mixed Damped Systems. If equation (2.1) is mixed damped, then at least one of the functions $a_n(t)$ in the series solution of (2.3) is overdamped and at least one is underdamped.

This result again follows directly from equation (3.6). If $L_1^2 - 4L_2$ is indefinite then it has at least one eigenvalue that is positive and at least one that is negative. It has been shown in the foregoing that the eigenvalues of $L_1^2 - 4L_2$ are given by (3.6), the discriminant that determines the oscillatory nature of $a_n(t)$. It follows that there will be at least one $a_n(t)$ that oscillates and at least one $a_n(t)$ that does not.

It should be noted that all the preceding results are actually both necessary and sufficient. This follows from the assumptions that L_1 and L_2 commute and each has a compact resolvent. The partial differential equation of motion (2.1) then reduces to an infinite set of decoupled ordinary differential equations. The discriminant of each of these equations is given by (3.6). Thus, if each $a_n(t)$ is overdamped (critically damped, underdamped, mixed), then each of the expressions (3.6) must be positive (zero, negative, of mixed signs). But (3.6) also expresses the eigenvalues of the operator $L_1^2 - 4L_2$; hence each of the eigenvalues are positive (zero, negative, of mixed signs). This means that $L_1^2 - 4L_2$ is positive definite (zero, negative definite, indefinite). Thus all of the results are necessary as well as sufficient.

4 Examples

Several examples serve to illustrate the validity of the results and to indicate their use.

Example 1. Consider the transverse free vibration of a membrane in a surrounding medium furnishing resistance to the motion that is proportional to the velocity. The equation of motion is

$$u_{tt}(\underline{x},t) + 2\gamma\, u_t(\underline{x},t) - \nabla^2 u(\underline{x},t) = 0 \quad \text{on} \quad \Omega$$

with boundary conditions

$$u(\underline{x},t) = 0 \quad \text{on} \quad \partial\Omega,$$

where ∇^2 is the harmonic operator in the Cartesian coordinates $\underline{x} = (x_1, x_2)$ and $u(\underline{x}, t)$ is the transverse deflection of the membrane. The domain Ω is a bounded open region in R^2 with boundary $\partial\Omega$. The positive constant γ reflects the proportional resistance to motion.

Denoting the λ_n as the eigenvalues of the harmonic operator on the domain Ω, then the time solutions $a_n(t)$ are given in [8] as

$$a_n(t) = e^{-\gamma t}\frac{\cos[\lambda_n - \gamma^2]^{1/2} t}{(\lambda_n - \gamma^2)^{1/2}}$$

if $\lambda_n > \gamma^2$. Hence, if $\lambda_n > \gamma^2$ for all n, then all the time solutions are underdamped. We claim that this same conclusion may be obtained using the result of Section 3 in this paper.

We identify the following:

let $L_1 = 2\gamma$, $L_2 = -\nabla^2$. We define

$\mathfrak{D}(L) = \{u(\cdot,t) \in \mathfrak{L}^2(\Omega)$ such that $u_{x_i}(\cdot, t)$ and $u_{x_i x_j}(\cdot, t)$ are continuous on Ω and $u(\underline{x}, t) = 0$ on $\partial\Omega\}$[2].

An easy calculation shows that both operators are positive and self-adjoint on $\mathfrak{D}(L)$. Also, they commute on $\mathfrak{D}(L)$.

In the foregoing we see that the eigenvalues of the operator

$$4L_2 - L_1^2 = -4(\nabla^2 + \gamma^2)$$

are $4(\lambda_n - \gamma^2)$. If $\lambda_n > \gamma^2$ for all n, then the operator $4L_2 - L_1^2$ is positive definite and all of the time functions in the series solution are underdamped—precisely the result found in [8].

Example 2. Consider the longitudinal vibration of a clamped bar with both internal and external damping. Under the assumption of linearity we write the equation of motion as

$$u_{tt}(x,t) + 2\left[\gamma - b\frac{\partial^2}{\partial x^2}\right]u_t(x,t) - a\frac{\partial^2 u}{\partial x^2}(x,t) = 0$$

for $x \in (0, 1) = \Omega$

where $u(x, t)$ is the displacement of the bar and the positive constants γ, b, and a reflect the relevant physical parameters.

The boundary conditions are taken to be

$$u(0,t) = u(1,t) = 0.$$

Before applying the results of this paper, we consider a series solution of the form

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t)\phi_n(x)$$

where the set $\{\phi_n(x)\}_{n=1}^{\infty}$ represent the orthonormal set of eigenfunctions obtained from the "restricted" eigenvalue problem

$$-a\phi''(x) = \lambda\,\phi(x), \quad \phi(0) = \phi(1) = 0.$$

Obviously one obtains an infinite number of discrete eigenvalues, λ_n, each of which is positive.

Using the series representation and the orthonormality of $\{\phi_n(x)\}_{n=1}^{\infty}$, we obtain an infinite set of decoupled equations

$$\ddot{a}_m(t) + 2\left[\gamma + \frac{b}{a}\lambda_m\right]\dot{a}_m(t) + \lambda_m a_m(t) = 0, \quad m = 1,2,3, \ldots .$$

The discriminant for each of these equations is of the form

$$4\left[\left(\gamma + \frac{b}{a}\lambda_n\right)^2 - \lambda_n\right].$$

Thus the case for over, under, critical, and mixed damping can be made by appealing to the sign of the discriminant. We show that the same results can be obtained in the following way.

Let $L_1 = 2\left[\gamma - b\frac{\partial^2}{\partial x^2}\right]$, $L_2 = -a\frac{\partial^2}{\partial x^2}$ with

$\mathfrak{D}(L) = \{u(\cdot, t) \in \mathfrak{L}(\Omega)$ such that all partial derivatives up to order four are continuous on (0, 1) and $u(0, t) = u(1, t) = 0\}$.

Note that L_1, L_2 are positive definite and commute on $\mathfrak{D}(L)$. Also, we calculate

[2] Trivially the assumptions in the problem definition are met with these conditions.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$L_1^2 = 4\left(\gamma^2 - 2\gamma b\frac{\partial^2}{\partial x^2} + b^2\frac{\partial^4}{\partial x^4}\right)$$

and

$$L_1^2 - 4L_2 = 4\left(\gamma^2 - 2\gamma b\frac{\partial^2}{\partial x^2} + b^2\frac{\partial^4}{\partial x^4} + a\frac{\partial^2}{\partial x^2}\right).$$

The eigenvalues of this operator are

$$4\left(\gamma^2 + 2\gamma\frac{b}{a}\lambda_n + \frac{b^2}{a^2}\lambda_n^2 - \lambda_n\right) = 4\left[\left(\gamma + \frac{b}{a}\lambda_n\right)^2 - \lambda_n\right].$$

Demanding that the sign of this expression be positive, negative, or zero is equivalent to characterizing the operator $4L_2 - L_1^2$. The conclusions to be drawn concerning the time solutions found in the infinite series are precisely the same as those obtained from the discriminant.

Note that since $\lambda_n = an^2\pi^2$, the only possibilities for damping are either overdamping for each mode or mixed damping with all "higher" modes overdamped.

Example 3. As a final example, we consider the free flexural vibrations of a damped plate. Previously this problem has been considered by Murthy and Sherbourne [9].

The equation of motion is given by

$$u_{tt}(\underline{x},t) + \frac{k}{mh}u_t(\underline{x},t) + \frac{D}{a^4mh}\nabla^4 u(\underline{x},t) = 0 \quad \text{on} \quad \Omega$$

with boundary conditions

$$u(\underline{x},t) = 0, \quad \frac{\partial u}{\partial n}(\underline{x},t) = 0 \quad \text{on} \quad \partial\Omega$$

where ∇^4 is the biharmonic operator in Cartesian coordinates $\underline{x} = (x_1, x_2)$, m, h, a, k, and D are all positive constants related to the physical parameters, and $\partial/\partial n$ indicates the normal derivative.

We identify the problem at hand in the following manner. Let $L_1 = k/mh = \beta^{1/2}$ and $L_2 = \gamma\nabla^4$, where $\gamma = D/mha^4$.

We define $\mathfrak{D}(L) = \{u(\cdot, t) \in \mathfrak{L}^2(\Omega)$ such that all partial derivatives up to order four are continuous on Ω and $u(x, t) = 0$, $\partial u/\partial n\ (x, t) = 0$ on $\partial\Omega.\}$ An easy calculation verifies the positivity and commutivity assumptions, as well as those of self-adjointness.

We consider the operator $L_1^2 - 4L_2 = \beta - 4\gamma\nabla^4$ on $\mathfrak{D}(L)$ and note that its eigenvalues are $(\beta - 4\gamma\lambda_n)$ where λ_n are the eigenvalues of ∇^4 on $\mathfrak{D}(L)$. Obviously these eigenvalues will be positive or negative depending on the value of n and the physical parameters k, m, D, h, and a. Surely, for "large" values of n, the sign of $(\beta - 4\gamma\lambda_n)$ will ensure that the time functions in the series solution will be underdamped while for "small" values of n, an overdamped time function may be possible.

Discussion of Results

The foregoing results present concise tests that may be used to determine whether a system will oscillate in time. This same test may also be used as a design tool in choosing damping and stiffness parameters for a system of given structure or as constraints in an optimal control problem. For instance, the parameters a, b, and γ of Example 2 could be chosen in such a way as to produce an overdamped system. If a, b, and γ are chosen so that $2b\gamma = a$, then $L_1^2 - 4L_2$ is positive definite on $\mathfrak{D}(L)$ and the system will be overdamped. On the other hand, the inequality $\gamma + bn^2\pi^2 > a^{1/2}n\pi$ could be used as a constraint in an attempt to minimize energy loss (or some other cost) in a velocity feedback scheme [10].

These conditions also lend themselves to analyzing the inherent nature of certain structures. As an example, in [6] Beskos and Boley conclude after analyzing several curves that complete overdamping can be achieved only for infinite values of damping for certain structures (i.e., Example 2 with $b = 0$). This conclusion can be reached immediately from the theory presented here by examining the inequality resulting from $L_1^2 - 4L_2$ being positive definite on $\mathfrak{D}(L)$ for the case $b = 0$. This requires $\gamma^2 > \lambda_n$ for all values of n. Since $\lambda_n \to \infty$, γ must of course become infinite for overdamping to occur in every mode. In addition this result indicates exactly the number of modes that will be overdamped for a given value of γ. It should also be noted that in all the examples considered, the condition of complete critical damping cannot be achieved, since the damping coefficient is not a function of the index n.

Other conclusions can also be stated. For instance, Example 2 shows that an internally damped beam ($\gamma = 0$) is inherently overdamped in as much as the terms $a_n(t)$ become overdamped for large n. On the other hand an externally damped beam ($b = 0$) will always have underdamped terms for large n.

Finally it should be noted that the results for the lumped parameter case [5] were stated in terms of the operator (matrix) $C - 2K^{1/2}$. In the distributed case presented here the operator $L_1^2 - 4L_2$ was chosen instead of $L_1 - 2L_2^{1/2}$ because of the difficulty in calculating the square root of an unbounded operator.

References

1 Duffin, R. J., "A Minimax Theory for Overdamped Networks," *Journal of Rational Mechanics and Analysis*, Vol. 4, 1955, pp. 221–233.

2 Nicholson, D. W., "Comments on Damped Response in Linear Systems," *Mechanics Research Communications*, Vol. 6, No. 1, 1979, pp. 17–25.

3 Müller, P. C., "Oscillatory Damped Linear Systems," *Mechanics Research Communications*, Vol. 6, No. 2, 1979, pp. 81–85.

4 Beskos, D. E., and Boley, B. A., "Critical Damping in Linear Discrete Dynamics Systems," ASME JOURNAL OF APPIED MECHANICS, Vol. 47, 1980, pp. 627–630.

5 Inman, D. J., and Andry, A. N., "Some Results on the Nature of Eigenvalues of Discrete Damped Linear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 927–930.

6 Beskos, D. E., and Boley, B. A., "Critical Damping in Certain Linear Continuous Dynamic Systems," *International Journal of Solids and Structures*, Vol. 47, 1981, pp. 575–588.

7 Caughey, T. K., and O'Kelly, M. E. J., "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, 1965, pp. 583–588.

8 Stakgold, I., *Boundary Value Problems of Mathematical Physics*, Vol. 2, Macmillan, New York, 1967.

9 Murthy, D. N. S., and Sherbourne, A. N., "Free Flexural Vibrations of Damped Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, 1972, pp. 298–300.

10 Balas, M. J., "Direct Velocity Feedback Control of Large Space Structures," *Journal of Guidance and Control*, Vol. 2, No. 3, 1979, pp. 252–253.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Compact Probabilistic Representation of Random Processes

S. F. Masri
Professor.
Mem. ASME

R. K. Miller
Associate Professor.
Mem. ASME

Civil Engineering Department,
University of Southern California,
Los Angeles, Calif. 90007

A method is given for representing analytically defined or data-based covariance kernels of arbitrary random processes in a compact form that results in simplified, analytical, random-vibration transmission studies. The method uses two-dimensional orthogonal functions to represent the covariance kernel of the underlying random process. Such a representation leads to a relatively simple analytical expression for the covariance kernel of the linear system response which consists of two independent groups of terms: one reflecting the input characteristics, and the other accounting for the transmission properties of the excited dynamic system. The utility of the method is demonstrated by application to a covariance kernel widely used in random-vibration studies.

1 Introduction

A review of the current state of the art of probabilistic methods and their use in applied mechanics indicates a very active interest in this important field and a growing recognition by designers of the significance of probabilistic approaches in engineering applications, particularly in the areas of probabilistic mechanics and structural reliability. A case in point is the recent ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability [1] during which about 100 technical papers were presented, spanning a wide range of topics.

A major part of the impetus behind this strong interest in probabilistic methods is the need to quantify the margin of safety in critical engineering facilities that are subjected to stochastic dynamic environments. For example, assuring the reliability of structural/mechanical components in a nuclear power plant subjected to earthquake loads is a problem that can be best handled in a probabilistic format.

In performing a probabilistic analysis of an equipment/structure model under seismic loads, the designer is faced with the problem of choosing an appropriate form of the input [2]: (1) design spectra, modified in various ways by site conditions, etc.; (2) selected samples of recorded earthquakes normalized in some way; (3) artificial earthquake records designed to match certain average characteristics of real events; or (4) a probabilistic description of the input based on an analysis of recorded earthquake ground motions.

For an accurate analysis of the response of a dynamic system to earthquake excitation, it is imperative that the evolutionary nature of the seismic loads be properly represented as a nonstationary stochastic process. This fact has motivated a considerable amount of research into the simulation of, and response to, nonstationary random processes. (For appropriate bibliography, see the review paper by Vanmarcke [3].) While these simulation studies have made available useful tools for the modeling of earthquakes in numerical computer studies, their format is not suitable for incorporation into an *analytical* random-vibration analysis. The difficulty lies in the fact that the approximations that have been developed for the simulation procedures often lead to a characterization of the earthquake loads so complicated that analytical response calculations become unduly cumbersome.

A popular procedure for analytically representing earthquake ground-motion time histories is to model them as zero-mean white noise processes modulated by a deterministic envelope function. While this approach is a step in the right direction, it suffers from the deficiency of not being an authentic representation of actual earthquake spectral contents. On the other hand, models of earthquakes as filtered white noise, where the filter has constant or time-varying parameters, lead to analytical expressions that cannot be conveniently handled in analytical, nonstationary, random-vibration solutions of the covariance kernels of dynamic systems response.

With the preceding discussion in mind, this paper presents a procedure for the compact probabilistic representation of nonstationary random phenomena, while at the same time allowing the use of the resulting characterization in analytical random-vibration studies.

2 Formulation

Consider a nonstationary random process $\{x(t)\}$ for which a statistically significant number of member function records $x_i(t)$ are available. The covariance kernel for $\{x(t)\}$ can be found from

$$K_{xx}(t_1,t_2)=E\{[x(t_1)-\mu_x(t_1)][x(t_2)-\mu_x(t_2)]\} \quad (1)$$

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October, 1981; final revision, March 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

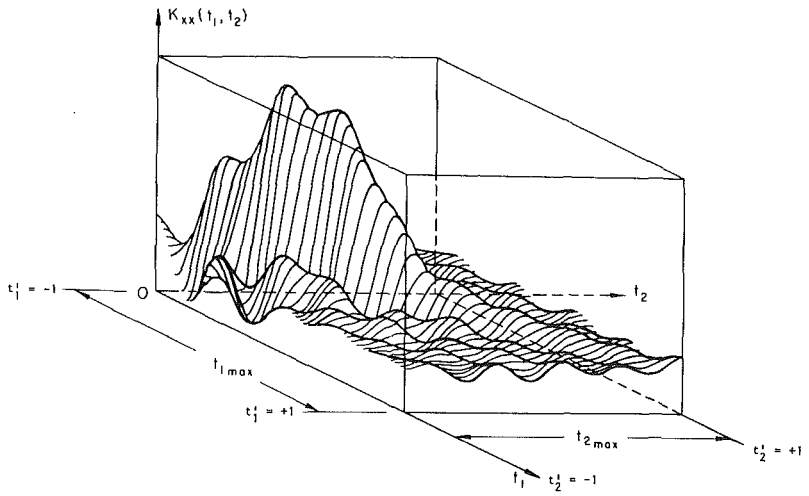


Fig. 1 Typical covariance function surface for a nonstationary process

where $E\{\ \}$ denotes the ensemble averaged value of the quantity within the braces and $\mu_x(t) = E[x(t)]$.

An example covariance function $K_{xx}(t_1,t_2)$ over a time domain $0 \le t_1 \le t_{1\max}$ and $0 \le t_2 \le t_{2\max}$ is shown in Fig. 1. While there are many methods available to graphically depict functions or collections of data in a three-dimensional form, a very limited subset of these methods yields analytical representations that can be conveniently incorporated into analytical probabilistic structural dynamic-response calculations.

A suitable candidate for the choice of functions to be used in the two-dimensional least-squares fit for the surface associated with $K_{xx}(t_1,t_2)$ are Chebyshev polynomials which, in addition to their orthogonal nature, have the desirable feature of equal-error approximation within an interval of interest.

2.1 Procedure for Determining Surface Fit. The main idea of the present procedure is to estimate the actual covariance kernel $K_{xx}(t_1,t_2)$ by an approximate function $\hat{K}_{xx}(t_1,t_2)$ expressed in terms of two-dimensional orthogonal polynomials. Thus,

$$K_{xx}(t_1,t_2) \simeq \hat{K}_{xx}(t_1,t_2) = \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij}T_i(t_1')T_j(t_2') \tag{2}$$

where the normalized values t_1' and t_2' of t_1 and t_2 are defined by

$$t_k' = 2(t_k/t_{k\max}) - 1, \quad k = 1,2 \tag{3}$$

and the T's are Chebyshev polynomials.

The least-squares Chebyshev polynomial approximation $\hat{K}_{xx}(t_1,t_2)$ of $K_{xx}(t_1,t_2)$ results in the coefficients C_{ij} of equation (2) being given by

$$C_{ij} = \begin{cases} (2/\pi)^2 v_{ij} & i \text{ and } j \neq 0 \\ (2/\pi^2) v_{ij} & i \text{ or } j = 0 \\ (1/\pi^2) v_{ij} & i = j = 0 \end{cases} \tag{4}$$

where

$$v_{ij} = \int_{\theta=0}^{\pi}\int_{\varphi=0}^{\pi} K_{xx}(\cos^{-1}\xi_1, \cos^{-1}\xi_2)T_i(\theta)T_j(\varphi)d\theta d\varphi \tag{5}$$

with $\xi_1 = \cos\theta$ and $\xi_2 = \cos\varphi$

2.2 Nonstationary Response of a SDOF System. Consider the nonstationary response $y(t)$ of a linear single degree-of-freedom (SDOF) system, characterized by an impulse function $h(t)$, under the action of a nonstationary stochastic process $\{x(t)\}$ whose covariance kernel is expressed as

$$\hat{K}_{xx}(t_1,t_2) = \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij}T_i(\xi_1)T_j(\xi_2) \tag{6}$$

where

$$\xi_i = \frac{2t_i}{T_m} - 1, \quad -1 \le \xi_i \le 1, \quad i = 1,2$$

and T_m = duration of the random excitation record.

For convenience, assume that the input is a zero-mean process. Then the covariance of the system output is

$$K_{yy}(t_1,t_2) = E[y(t_1)y(t_2)] \tag{7}$$

with

$$E[y(t_1)y(t_2)] = \int_0^{t_1}\int_0^{t_2} h(t_1-\tau_1)h(t_2-\tau_2)K_{xx}(\tau_1,\tau_2)d\tau_2 d\tau_1 \tag{8}$$

Replacing K_{xx} in equation (8) by its representation in equation (6), the covariance of the response will then be given by

$$E[y(t_1)y(t_2)] = \frac{T_m^2}{4}\sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij}I_i(t_1)I_j(t_2) \tag{9}$$

where

$$I_i(t) = \int_{-1}^{\left(\frac{2t}{T_m}-1\right)} h\left(t - \frac{T_m}{2}(\xi+1)\right)T_i(\xi)d\xi \tag{10}$$

This integral may be evaluated by expressing the $T_i(\xi)$ as an ith order polynomial in ξ:

$$T_i(\xi) = \sum_{k=0}^{i} f_{ik}\xi^k \tag{11}$$

Consider a lightly damped linear SDOF oscillator with an impulse function

$$h(t) = -\frac{1}{\omega_d}e^{-\zeta\omega_n t}\sin\omega_d t \tag{12}$$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

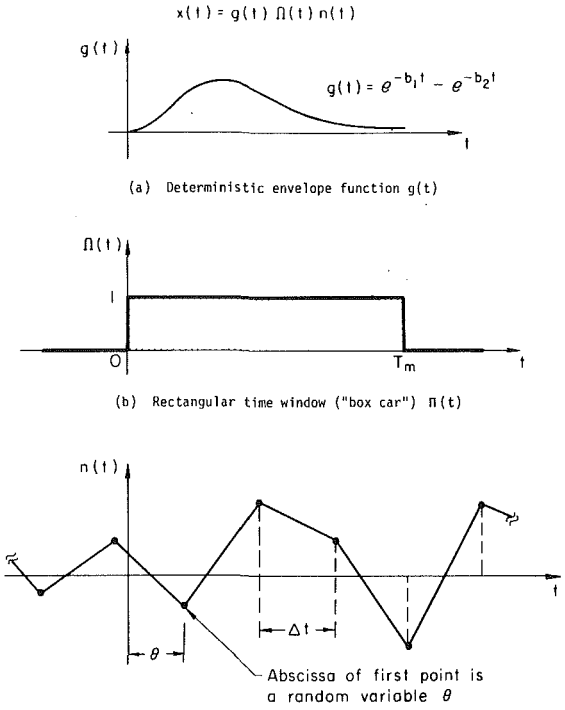


Fig. 2 Characteristics of example random process

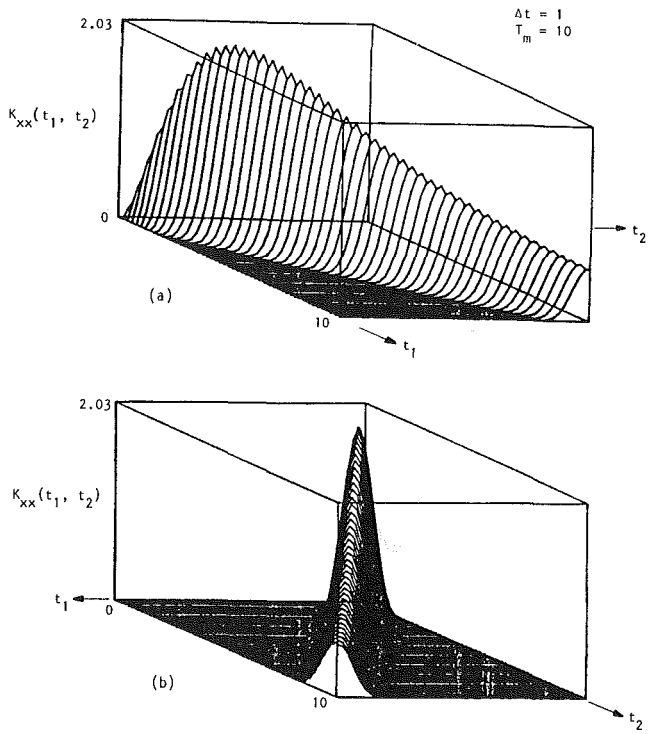


Fig. 3 Exact covariance kernel $K_{xx}(t_1,t_2)$

where

ζ = ratio of critical damping
ω_n = natural frequency
$\omega_d = \omega_n\sqrt{1-\zeta^2}$

Making use of equations (11) and (12) the value of the generic integral appearing in equation (10) may be written explicitly as

$$I_i(t) = -\frac{1}{\omega_d} e^{(a-\zeta\omega_n t)} F_i(t) \tag{13}$$

where

$$F_i(t) \equiv \left[e^{a\xi} \sum_{k=0}^{i} \sum_{l=k}^{i} \frac{(-1)^k}{\rho^{k+1}} \frac{l!}{(l-k)!} f_{il} \right.$$
$$\left. \xi^{l-k} \sin(b\xi + \beta + (k+1)\alpha) \right]_{\xi=-1}^{\left(\frac{2t}{T_m} - 1\right)} \tag{14}$$

and where

$$a = \frac{\zeta\omega_n T_m}{2}; \quad b = -\frac{\omega_d T_m}{2}; \quad \rho = \omega_n T_m/2;$$
$$\alpha = \tan^{-1}\left(\sqrt{1/\zeta^2 - 1}\right); \quad \beta = \beta(t) = \omega_d\left(t - \frac{T_m}{2}\right)$$

Using (13) in (9), the covariance of the response is

$$E[y(t_1)y(t_2)] = \sum_{i=0}^{m} \sum_{j=0}^{n} C_{ij} Y_{ij}(t_1,t_2) \tag{15}$$

where

$$Y_{ij}(t_1,t_2) \equiv \frac{T_m^2}{4\omega_d^2} \exp[2a - \zeta\omega_n(t_1+t_2)] F_i(t_1) F_j(t_2) \tag{16}$$

2.3 Computational Considerations. From the definition of the excitation covariance kernel, $K_{xx}(t_1,t_2)$ is a symmetric function of its arguments t_1 and t_2; consequently the matrix $[C]$ of the coefficients associated with the approximate surface $\hat{K}_{xx}(t_1,t_2)$ is symmetric and only the upper or lower triangular portion of $[C]$ need be determined.

Also, most physical random proccesses of the type usually encountered in applied mechanics problems will have a K_{xx} surface whose peaks (mean-square values) lie along the diagonal of the (t_1,t_2) plane in the form of a "ridge" whose width (in direction $\tau = t_1 - t_2$ orthogonal to the main diagonal) is relatively narrow. In other words, the time period $\tau = t_1 - t_2$ during which the autocorrelation function of the excitation at any time t_1 reduces from its peak value to a negligible amount is relatively short (i.e., the sides of the K_{xx} ridge have steep slopes). These qualitative features of the covariance kernel cause the matrix $[C]$ to assume a banded form which further reduces the number of C_{ij}'s to be determined in practical applications.

3 Applications

Consider a random process $x(t)$ defined by

$$x(t) = g(t)\Pi(t)n(t) \tag{17}$$

where $g(t)$ is a deterministic envelope function

$$g(t) = \exp(-b_1 t) - \exp(-b_2 t) \tag{18}$$

in which b_1 and b_2 are positive constants; $\Pi(t)$ is a rectangular window ("box car") function defined as

$$\Pi(t) = \begin{cases} 0 & ; \ 0 > t \\ 1 & ; \ 0 \le t \le T_m \\ 0 & ; \ t > T_m \end{cases} \tag{19}$$

and $n(t)$ is a random process in which a sample function is created by assigning statistically independent values of n to successive ordinates equally spaced along the time axis. It is assumed that values of $n(t)$ for any time can be found by linear interpolation, as shown in Fig. 2. The abscissa of the first point in the range $t > 0$ is a random variable θ that has a

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

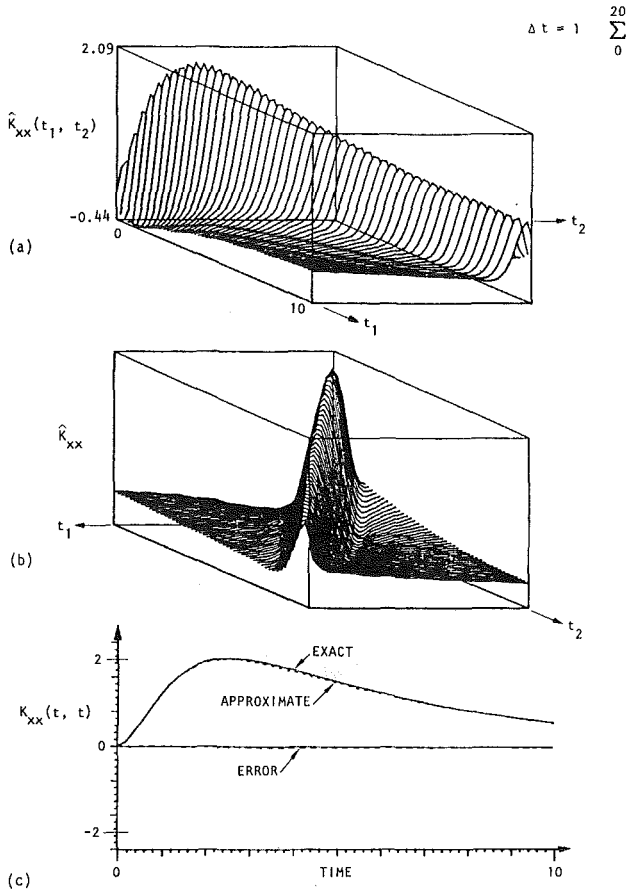


Fig. 4 Approximate covariance kernel $\hat{K}_{xx}(t_1,t_2)$, $i_{max} = j_{max} = 20$

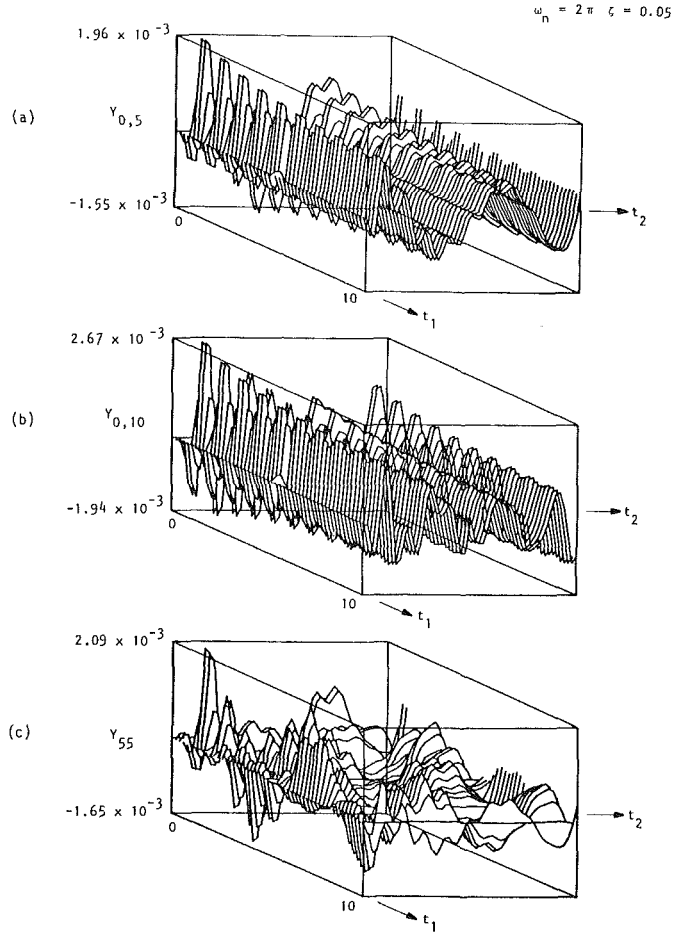


Fig. 5 Three-dimensional representation of typical $Y_{ij}(t_1,t_2)$.

uniform probability density function of magnitude $1/\Delta t$ over the time interval $0<t<\Delta t$, where Δt is the constant time spacing of the ordinates shown in Fig. 2. Let the random process $n(t)$ have a zero mean and a Gaussian probability density function with variance σ_0^2, given by

$$\sigma_0^2 = 2\pi S_0/(\Delta t) \tag{20}$$

where S_0 is a constant. Then the autocorrelation function $R_n(\tau)$ for the random process $n(t)$ is given by [4]

$$R_n(\tau) = \begin{cases} \sigma_0^2\left[\frac{2}{3} - \left(\frac{\tau}{\Delta t}\right)^2 + \frac{1}{2}\left(\frac{|\tau|}{\Delta t}\right)^3\right], & |\tau| < \Delta t \\ \sigma_0^2\left[\frac{4}{3} - 2\frac{|\tau|}{\Delta t} + \left(\frac{\tau}{\Delta t}\right)^2 - \frac{1}{6}\left(\frac{|\tau|}{\Delta \tau}\right)^3\right], & \Delta t \le |\tau| \le 2\Delta t \\ 0, & 2\Delta t < |\tau| \end{cases} \tag{21}$$

where $\tau = t_2 - t_1$ and the spectral density function of $n(t)$ is

$$S_n(\omega) = S_0 \frac{6 - 8\cos(\omega\Delta t) + 2\cos(2\omega\Delta t)}{(\omega\Delta t)^4}, \quad -\infty < \omega < \infty \tag{22}$$

It is worth noting from (17) and (21) that as $\Delta t \to 0$,

$$\lim_{\Delta t \to 0} K_{xx}(t_1,t_2) = g(t_1)g(t_2)\Pi(t_1)\Pi(t_2)2\pi S_0\delta(t_2 - t_1) \tag{23}$$

where $\delta(\tau)$ is the Dirac delta function.

3.1 Chebyshev Fit for $K_{xx}(t_1,t_2)$.

Case 1. Consider an example random process of the type defined by equation (17) with the following parameters:

$$b_1 = 1, \quad b_2 = 0.1, \quad T_m = 10, \quad \Delta t = 1, \quad S_0 = 1 \tag{24}$$

A three-dimensional plot of this covariance function $K_{xx}(t_1,t_2)$ with both t_1 and t_2 varying between 0 and T_m is shown in Fig. 3. Note that the peak value of K_{xx} is ≈ 2 and it occurs at $t_1 = t_2 \approx 2.5$. Figure 3(*b*) represents the surface as viewed from a position looking directly into the diagonal $t_1 = t_2 = t$ at the end corresponding to $t = T_m$.

A Chebyshev polynomial fit for this exact function K_{xx} is shown in Fig. 4 where parts (*a*) and (*b*) show two views corresponding to those of Fig. 3 of the surface associated with the approximate kernel $\hat{K}_{xx}$, and part (*c*) shows a comparison between the exact and approximate kernel along the diagonal of the t_1,t_2 plane. The maximum number of terms used in the series expansion of equation (6) along each of the time axes (i_{max} for the t_1 axis and j_{max} for the t_2 axis) are both equal to 20.

The values of the C_{ij} coefficients for this case clearly show the dominance of the diagonal terms and the fact that entries outside a band of five elements on either side of the diagonal are negligible. This feature of matrix [*C*] results in a substantial reduction of the number of terms to be used for a given order n from n^2 in the case of a general matrix [*C*] to $n(1+N_c)$ in the case where a band of $\pm N_c$ terms are to be used about the diagonal of symmetric matrix [*C*].

Case 2. This case is identical to the preceding one except that the value of Δt is reduced by an order of magnitude from 1.0 to 0.1. The three-dimensional view of K_{xx} in a form

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

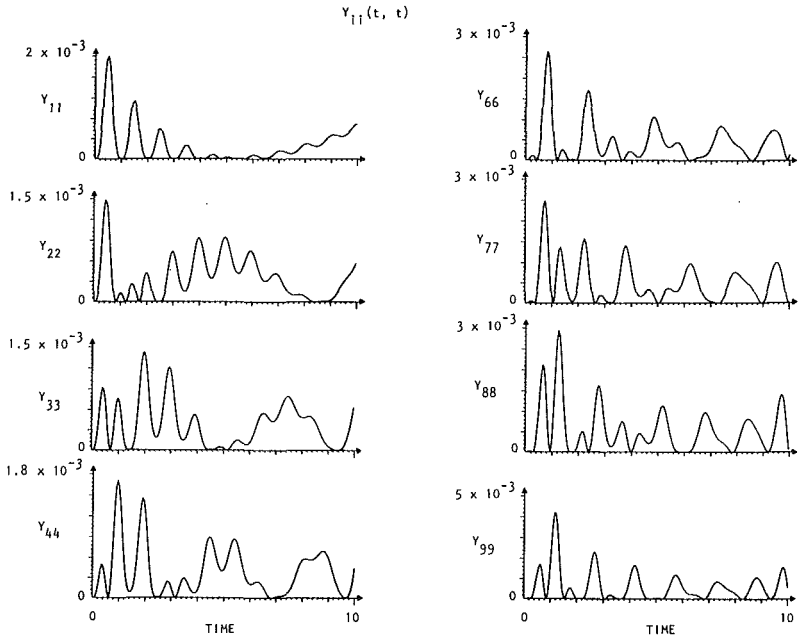


Fig. 6 Time histories of sample $Y_{ii}(t,t)$ functions

similar to that of Fig. 3(*b*) will result in essentially a "knife edge" shape due to the extreme sharpness of the surface variation in the neighborhood of the diagonal.

To achieve a degree of accuracy in the fit between the exact and approximate corvariance surfaces comparable to the graphical results shown in Fig. 4(*c*) required about 100 diagonal terms with a band $N_c \approx 15$. This behavior is consistent with the variation of the number of terms used in a series of a single-variable function with the severity of rapid change in the function.

3.2 Transient Mean-Square Response. Consider a linear single-degree-of-freedom harmonic oscillator with a natural frequency $\omega_n = 2\pi$ and a ratio of critical damping $\zeta = 0.05$. Using equation (16), values of representative $Y_{ij}(t_1,t_2)$ functions associated with this oscillator are shown in Figs. 5 and 6.

Note from the definition of $Y_{ij}(t_1,t_2)$ that:

(*a*) the Y_{ij}'s are symmetric with respect to their arguments: $Y_{ij}(t_1,t_2) = Y_{ji}(t_2,t_1)$;

(*b*) $Y_{ii}(t,t) \geq 0$, $i = 0,1,2, \ldots$;

(*c*) $Y_{ij}(t,t)$ may assume negative, zero, or positive values, depending on the value of t and system parameters;

(*d*) The time of occurrence of various peaks of Y_{ij} may assume any value in the range $0 \leq t \leq T_m$;

(*e*) the predominant frequency of variation in the magnitude of Y_{ij} corresponds to twice the oscillator natural frequency (note the presence in equation (15) of terms involving $\sin^2(\omega_d t)$ which will give rise to $\sin(2\omega_d t)$ terms);

(*f*) the attenuation in the amplitude of oscillations of a particular Y_{ij} is related to the extent of damping present in the system.

Case 1. Consider now the transient mean-square response of the example oscillator under the action of the nonstationary excitation characterized by the covariance kernel shown in Fig. 3. Using the appropriate C_{ij} coefficients and the corresponding Y_{ij}'s as shown in equation (15) results in the response shown in Fig. 7.

The results of Fig. 7 compare the accuracy of various approximate solutions as the number of terms used in equation (15) is changed from $m = n = 0$ to 5,10,15, and 20. The exact solution shown in Fig. 7 was obtained by numerically evaluating equation (8) with the exact value of K_{xx} as defined by equations (17)–(23).

In each of the five plots shown in Fig. 7, the solid line represents the exact mean-square response $E[y^2(t)]$ and the dashed line represents the approximate mean-square response $E[\hat{y}^2(t)]$ achieved by summing all the terms in the double-indexed series up to the term indicated in the figure, and the broken line represents the contribution of the last terms, $\Delta E[\hat{y}^2(t)]$, added to the approximate mean-square solution, i.e., the terms of the double series associated with the indicated highest index of the double series.

Case 2. For this case, the K_{xx} associated with the smaller $\Delta t = 0.1$ is applied to the same oscillator discussed in the foregoing. Repeating the steps of the previous example, mean-square response results comparable in accuracy to those of Fig. 7(*e*) are achieved by using 35 diagonal terms and a bandwidth of 15.

It is worth noting that the increase in steepness of the covariance function surface due to changing parameter Δt from 1.0 to 0.1 necessitated the use of about 5 times more diagonal terms and a bandwidth about 3 times wider than the first case in which $\Delta t = 1$, to obtain a comparable goodness-of-fit between *excitation kernel* K_{xx} and $\hat{K}_{xx}$. However, to attain comparable accuracy in the mean-square *response* calculations of $E[y^2(t)]$, the number of needed diagonal terms increased only by about a factor of two and the bandwidth again by about a factor of 3. This behavior is due to the filtering effects inherent in the oscillator dynamic characteristics.

4 Discussion

It is clear from the preceding discussion that the method presented in this paper provides a means for the compact representation of random processes in a probabilistic format that can be conveniently employed in random-vibration studies. Not only does this method provide a simple approximation for analytically defined, complicated covariance kernels which could not otherwise be conveniently used in analytical random vibration studies, but, more importantly, it

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

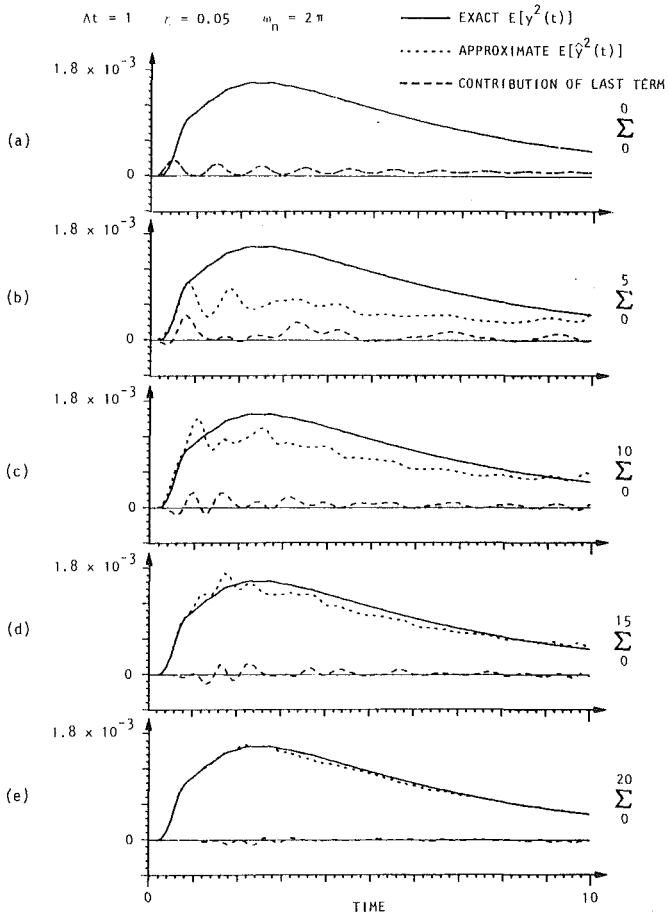


Fig. 7 Comparison of exact and approximate mean-square response

also provides a simple way to derive a data-based probabilistic description of random processes such as earthquake ground motion, wind loads, turbulence phenomena, etc., which otherwise would not be amenable to a compact probabilistic description. The method of this paper condenses the representation of random processes to the calculation of about one-half the entries in a symmetric matrix of constant coefficients.

The approximate mean-square response calculations are completely independent of the steps involved in processing the excitation. Once the input is characterized by means of the C_{ij}'s, no further modifications in these calculations are needed when investigating the response of different oscillators. The effects of the excitation on various oscillators can be evaluated by changing the Y_{ij} terms which depend only on the oscillator characteristics. Thus, random-vibration response calculations are reduced to the evaluation of a collection of simple algebraic expressions, whose evaluation does not require sophisticated techniques, and which can be easily calculated without resorting to any further approximations (e.g., no numerical integration is needed).

An important feature of the analysis technique is that the only approximation introduced in the mean-square response calculations involves the representation of the exact covariance kernel K_{xx} by the approximate $\hat{K}_{xx}$; no subsequent approximations have been made in deriving the expressions for the mean-square response. Furthermore, the formulation made no restrictive assumption concerning the nature of the convariance kernel so that completely general excitations may be considered.

An example involving an analytically defined excitation was used instead of a numerically defined one in order to examine in a straightforward manner the important considerations of the number of terms necessary to adequately model the excitation response as a function of the bandwidth of the excitation (or the digitization rate, Δt). Further details regarding this study are available in [5].

5 Summary and Conclusions

A relatively simple and straightforward procedure is given for representing analytically defined or data-based covariance kernels of arbitrary random processes in a compact form that allows its convenient use in later analytical random-vibration response studies.

The method is based on the use of suitable two-dimensional orthogonal functions to match the surface of the covariance kernel function $K_{xx}(t_1,t_2)$ covering the (t_1,t_2) plane. The form of the estimated surface $\hat{K}_{xx}(t_1,t_2)$ results in the mean-square reponse being given by a series of terms involving a product of two groups: (*a*) a constant coefficient that depends solely on the nature of the excitation, and (*b*) a time function expressed in an algebraic form that depends only on the oscillator characteristics.

The utility of the method for representing the excitation and calculating the mean-square response is illustrated by the use of an analytical expression for the covariance kernel associated with a widely used approximation to digitally generated, modulated, white-noise excitation functions.

Acknowledgments

This study was supported in part by a grant from the National Science Foundation. The assistance of T. K. Caughey, D. E. Hudson, and T. Dehghanyar in the various phases of this study is appreciated.

References

1 Ang, A. H-S., and Shinozuka, M., "Probabilistic Mechanics and Structural Reliability," *Proc. ASCE Conference*, Tucson, Ariz., Jan. 10–12, 1979.

2 Masri, S. F., Richardson, J. E., and Young, G. A. (1978), "Evaluation of Seismic Analysis Techniques for Nuclear Power Plant Piping and Equipment," presented at the Engineering Design for Earthquake Environments Conference, The Institute of Mechanical Engineers, Applied Mechanics Group, London, *Proc. I. Mech. E.*, Nov. 7–9, 1978, pp. 75–90.

3 Vanmarcke, E. H., "Some Recent Developments in Random Vibration," *Applied Mechanics Review*, Vol. 32, No. 10, 1979, pp. 1197–1202.

4 Clough, R. W., and Penzien, J., *Dynamics of Structures*, McGraw-Hill, New York, 1975.

5 Masri, S. F., and Miller, R. K., "Compact Probabilistic Representation of Random Processes," Univ. of Southern Calif. Civil Engineering Report No. USC-CE-8202, 1982.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

S. F. Masri
Professor.
Mem. ASME

G. A. Bekey
Professor.

School of Engineering,
University of Southern California,
Los Angeles, Calif. 90007

T. K. Caughey
Professor,
Division of Engineering and
Applied Science,
California Institute of Technology,
Pasadena, Calif. 91109

On-Line Control of Nonlinear Flexible Structures

A simple yet efficient method is presented for the on-line control of nonlinear, multidegree-of-freedom systems responding to arbitrary dynamic environments. The control procedure uses pulse generators located at selected positions throughout a given system. The degree of system oscillation near each controller determines the controller's activation time and pulse amplitude. The direct method of Liapunov is used to establish that the response of the controlled nonlinear system is Lagrange stable. Simulation studies of three example systems (conducted with digital and analog computers) demonstrate the feasibility, reliability, and robustness of the proposed active-control method. These systems, which include one with a hysteretic nonlinearity, are structures representative of modern tall buildings; they are subjected to nonstationary random excitation representative of earthquake ground motions.

1 Introduction

The active control of flexible systems is a subject of growing interest in the field of structural dynamics. References [1–7] are representative publications on the subject. The concept of active control of massive structural systems has received considerable attention. However, only a fraction of the published studies offers methods suitable for on-line control of typical flexible structural systems subjected to arbitrary dynamic environments. Most of these methods require either unreasonable computational effort for their complex feedback control laws or excessive control energy or both.

Recently, the authors presented a simple yet efficient active-control method for reducing the oscillations of flexible systems subjected to arbitrary deterministic or stochastic excitations [8]. This was an optimum pulse control method shown in the referenced paper to be suitable for on-line implementation to control the response of realistic structural systems. It requires a continuous monitoring of the systems state variables. When some specified threshold has been exceeded, an open-loop control pulse is applied. The analytical procedure for optimizing the needed pulse amplitudes of the various pulse generators (actuators) assumes that the pulsers, which are distributed in the structure, must be activated simultaneously and must have the same pulse duration. A least-squares optimization procedure is used that involves the modal analysis approach. Therefore, knowledge of the eigenvalues and eigenvectors of a linearized model of the structure to be controlled is needed at the onset of the active-control process.

In reality, active-control techniques will be applied to dynamic systems undergoing large deformations (for example, buildings subjected to strong earthquake ground motion). These systems will operate in their nonlinear range of motion. Very likely they will operate as deteriorating nonlinear systems with time-varying properties. Consequently, the application of the optimization scheme to arbitrary nonlinear systems with unknown characteristics would require the use of suitable system-identification techniques to determine an equivalent linearized model of the structure every time a control pulse was needed. Depending on the nature and order of the system to be controlled, this identification task might require a certain amount of computational effort (computer CPU time) which, when added to the computations needed for the optimization procedure, could preclude the on-line implementation of the algorithm.

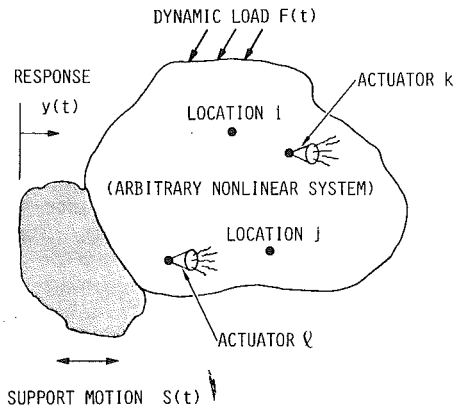


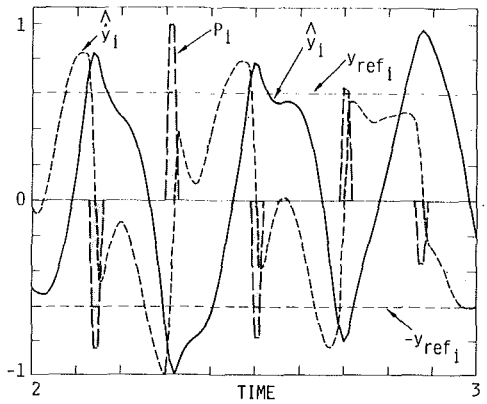
Fig. 1 Model of arbitrary, nonlinear distributed parameter system to be actively controlled

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

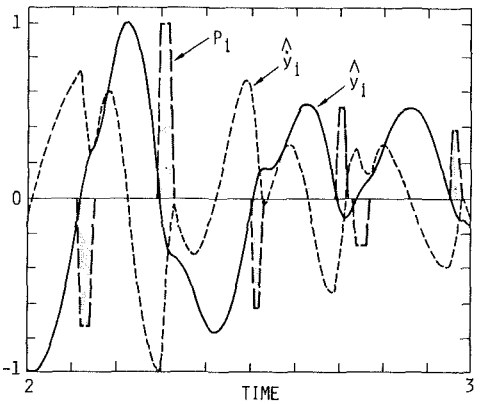
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May, 1981; final revision, March, 1982. Paper No. 82-WA/APM-25.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



(a) Threshold crossing



(b) Zero crossing

Fig. 2 Active control strategy

This paper presents another very simple and reliable method for on-line active control of the response of general, nonlinear, flexible structures (including those with time-varying properties) subjected to arbitrary dynamic excitations. It is shown by simulation studies that the proposed method is an efficient procedure for attenuating the response of realistic building-like structures under earthquake loads. It requires virtually no computational effort. In addition, the relatively modest amount of control energy required is shown to be within the capacity of portable gas-powered pulse generators currently under development [9].

2 Control Procedure

Consider the arbitrary, nonlinear, multidegree-of-freedom system shown in Fig. 1 under dynamic load $F(t)$, with support motion $S(t)$. The motion of certain locations in this structure is to be attenuated.

In the context of earthquake engineering, the objective is to lessen the potential damage to the building structure which would result from excessive deformations $y_i(t)$ relative to the foundation undergoing an earthquake ground acceleration $\ddot{S}(t)$. One approach to this problem is to disrupt the orderly buildup of amplitude associated with resonance. If pulses of suitable magnitude and proper direction are applied at several locations throughout the structure, resonance will be interrupted. The amount of control energy used can be minimized if the control action is activated only when the structural response exceeds a certain threshold related to the resistance of the structure.

Figure 2(*a*) illustrates a representative segment of the motion being controlled by optimum pulses triggered when the structural response exceeds some bounds [8]. This graph shows the relative displacement $\hat{y}_i(t)$, the relative velocity $\dot{\hat{y}}(t)$, the control force $p_i(t)$, and the displacement trigger levels. The time histories in Fig. 2(*a*) have been normalized to lie between $+1.0$ and -1.0.

As seen in Fig. 2(*a*), a suitable pulse is applied about once every fundamental period of the system. The optimum value of each pulse would depend on the state variables $\mathbf{y}(t_0)$ and $\dot{\mathbf{y}}(t_0)$ of the system at time t_0 when the barrier level $\mathbf{y}_{ref}$ was breached. Consequently, the amplitude of the pulses is proportional to the energy of the structure at t_0, which indirectly reflects the effects of the excitation energy from the time motion begins until time t_0. Note that each of the optimized pulses is always opposite in direction to the corresponding velocity at time t_0. Their magnitude is proportional to the instantaneous values of $\dot{\hat{y}}(t_0)$ at the corresponding threshold crossing.

The work done by a force p acting on a dynamic system with a velocity v is of the form

$$w = \int p(t)\, v(t)\, dt \tag{1}$$

This expression coupled with the response characteristics of Fig. 2(*a*), makes it clear that the most beneficial damping effects will be obtained from a control action that applies an opposing force while the system velocity is maximum. This instant corresponds to the zero crossings of the associated displacement and acceleration.

The active-control procedure discussed in this paper consists of using, at selected points in the structure, force actuators whose operation is triggered every time a zero crossing of the relative displacement of each of the points of interest is detected. The magnitude of each control force p_i will depend on the instantaneous magnitude of the corresponding velocity v_i and it will oppose v_i. Thus

$$p_i(t) = \begin{cases} -c_i \operatorname{sgn}(v_i)\, v_i^{n_i} & t_{0_i} \le t \le (t_{0_i} + T_{d_i}) \\ 0 & t > t_{0_i} + T_{d_i} \end{cases} \tag{2}$$

where

- c_i is a coefficient for scaling the needed control force at location i;
- sgn(•) indicates the sign of its argument;
- v_i is the absolute or relative velocity of the structure at location i, depending on the nature of the problem;
- n_i is some appropriate power of the velocity at location i;
- t_{0_i} is the zero crossing time at location i;
- T_{d_i} is the time duration of the controlling force at location i.

The simple nonlinear expression of equation (2), which relates the needed control force to an arbitrary power of the associated system velocity, has some worthwhile features:

1. The sgn(•) function ensures that the sign of the control force has the correct phase with respect to the system velocity.

2. If n is selected as $n=0$, the control force acts as a Coulomb friction force with magnitude $\pm c_i$.

3. If n is chosen as $n=1$, the control force introduces active viscous damping with a coefficient c_i.

4. If the value of n is chosen to be greater than 1, nonlinear velocity damping is introduced.

5. To optimize the control-energy effects, the pulse generators are activated for a relatively short duration T_{d_i} (when velocity is at its peak level) after the displacement zero crossing at location i, which occurs at time t_{0_i}.

6. If T_{d_i} for a given pulse generator is set greater than

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

about half a fundamental period of the system, the actuator switching will be controlled by the velocity change in sign instead of T_{d_i}. Thus, the actuator will act essentially as a negative feedback device whose action introduces nonlinear velocity-proportional damping.

Figure 2(*b*) shows a representative segment of the motion of an arbitrary system being controlled in accordance with equation (2), using a linear power of the velocity ($n=1$). Note that the application of the control pulse, which has a duration about 1/25 of the fundamental period, substantially reduces the system velocity and hence the system response.

To limit excessive demands for peak control force during episodes of high-level response, an upper limit could be imposed. Similarly, to conserve control energy when the system response is within acceptable range, the controller would be kept inactive as long as the force required were less than some threshold level. Both of these conditions can be conveniently incorporated in the control algorithm by augmenting equation (2) with the following constraints:

$$P_i(t)=\begin{cases}\pm P_{\max_i} & |P_i|>P_{\max_i}\\ 0 & |P_i|<P_{\min_i}\end{cases} \tag{3}$$

A more general form of equation (2) is to relate the magnitude of the control force at location i to both the displacement x_i and velocity v_i at that location:

$$P_i(t)=m_i(a_i x_i+b_i v_i)+g_i(x_i,v_i) \tag{4}$$

where m_i is the mass associated with i, and g_i is some suitable nonlinear function. The advantage of the formulation in equation (4) is that only *local* measurements of displacement and velocity are needed to determine the control force to be applied at a specific location. Since no global information regarding the stiffness and damping characteristics of the system is needed, the control law of equation (4) offers a distinct advantage in dealing with realistic structural systems whose stiffness and damping parameters are difficult to identify accurately in practical cases.

3 Stability Analysis

Assume that the n-degree-of-freedom dynamic system to be controlled is nonlinear and governed by the differential equation

$$M\ddot{\mathbf{x}}+D\dot{\mathbf{x}}+K\mathbf{x}+\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})+\nabla F(\mathbf{x})=\mathbf{q}(t) \tag{5}$$

with the following properties:

1. M is a diagonal positive definite matrix = $\lceil\, m \,\rfloor$.
2. D is a symmetric, positive-semidefinite, small matrix such that

$$\mathbf{x}^T K\mathbf{x}>c\mathbf{x}^T D\mathbf{x} \tag{6}$$

where c is a positive constant, $0<c<1$.

3. K is a symmetric, positive-definite matrix.
4. $\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})$ represents nonconservative nonlinear forces and

$$|f_i(\mathbf{x},\dot{\mathbf{x}})|\le B_i \text{ with } i\in(1,n) \tag{7}$$

and B_i are bounded positive constants.

5. $\nabla F(\mathbf{x})$ represents the conservative nonlinear forces which are equal to the gradient of function F, and

$$F(\mathbf{x})>0,\ \mathbf{x}^T\nabla F(\mathbf{x})\ge\alpha F(\mathbf{x}) \tag{8}$$

with constant $\alpha>1$.

6. $\mathbf{q}(t)$ is a bounded excitation vector which satisfies the condition

$$|q_i(t)|\le Q_i \text{ with } i\in(1,n) \tag{9}$$

and Q_i are bounded constants.

Note from equation (7) that nonlinear damping forces of the type inherent in bilinear hysteretic elements or, for example, Coulomb-type dissipative forces can be handled within the restriction imposed on $\mathbf{f}$. Since the damping matrix D is small and required to only be positive-semidefinite, and since the presence of nonconservative nonlinear forces does not guarantee a bounded response in the presence of persistent disturbances (e.g., Coulomb-damped or bilinear hysteretic systems under harmonic excitation), the response of the system governed by equation (5) may be large or even unbounded.

Suppose that a control force vector $\mathbf{h}$ is created such that

$$\mathbf{h}(\dot{\mathbf{x}}+c\mathbf{x})=-2cM(\dot{\mathbf{x}}+c\mathbf{x})-\mathbf{g}(\dot{\mathbf{x}}+c\mathbf{x}) \tag{10}$$

where c is a positive constant and

$$(\dot{\mathbf{x}}+c\mathbf{x})^T g(\dot{\mathbf{x}}+c\mathbf{x})\ge 0. \tag{11}$$

For the single degree-of-freedom case, the condition expressed by equation (11) is that function $g(\bullet)$ behaves like a straight line through the origin so that $yg(y)>0$ for $y\neq 0$ and $g(0)=0$.

The controlled dynamic system of equation (5) is now governed by

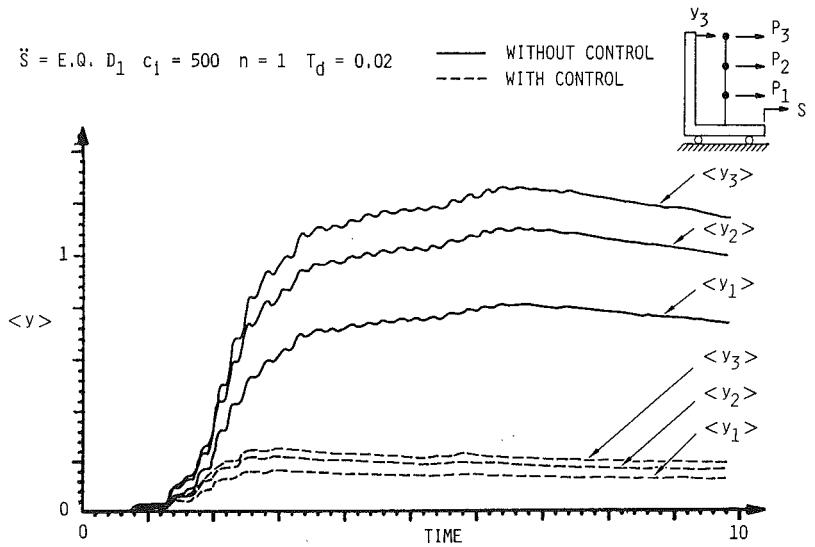


Fig. 3 Comparison of rms response with and without control

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$M\ddot{\mathbf{x}}+(D+2cM)\dot{\mathbf{x}}+(K+2c^2M)\mathbf{x}+\mathbf{f}(\mathbf{x},\dot{\mathbf{x}}) + \nabla F(\mathbf{x})+\mathbf{g}(\dot{\mathbf{x}}+c\mathbf{x})=\mathbf{q}(t) \quad (12)$$

It will now be shown that the solutions of equation (12) are Lagrange stable (bounded).

Using Liapunov's direct method, take the scalar function $V(\mathbf{x},\dot{\mathbf{x}},t)$ to be

$$V=\frac{1}{2}[(\dot{\mathbf{x}}+c\mathbf{x})^TM(\dot{\mathbf{x}}+c\mathbf{x})+\mathbf{x}^T(K+c^2M+cD)\mathbf{x}]+F(\mathbf{x}) \quad (13)$$

It is clear from equation (13) that V is a positive-definite function. The time derivative of V is

$$\dot{V}=(\dot{\mathbf{x}}+c\mathbf{x})^TM\ddot{\mathbf{x}}+(\dot{\mathbf{x}}+c\mathbf{x})^TcM\dot{\mathbf{x}} +\mathbf{x}^T(K+c^2M+cD)\dot{\mathbf{x}}+\dot{\mathbf{x}}^T\nabla F \quad (14)$$

Using equation (12),

$$\dot{V}=-\{\dot{\mathbf{x}}^TD\dot{\mathbf{x}}+c[(\dot{\mathbf{x}}+c\mathbf{x})^TM(\dot{\mathbf{x}}+c\mathbf{x})+\mathbf{x}^T(K+c^2M)\mathbf{x} +\mathbf{x}^T\nabla F(\mathbf{x})]\}+(\dot{\mathbf{x}}+c\mathbf{x})^T[\mathbf{q}(t)-\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})-\mathbf{g}(\dot{\mathbf{x}}+c\mathbf{x})] \quad (15)$$

Completing the squares, equation (15) can be written as

$$\dot{V}=-\left\{\dot{\mathbf{x}}^TD\dot{\mathbf{x}}+\frac{c}{2}(\dot{\mathbf{x}}+c\mathbf{x})^TM(\dot{\mathbf{x}}+c\mathbf{x}) +c\mathbf{x}^T(K+c^2M)\mathbf{x}+c\mathbf{x}^T\nabla F(\mathbf{x})\right\} -\frac{c}{2}\left[M^{1/2}(\dot{\mathbf{x}}+c\mathbf{x})-\frac{1}{c}M^{-1/2}(\mathbf{q}-\mathbf{f})\right]^T \left[M^{1/2}(\dot{\mathbf{x}}+c\mathbf{x})-\frac{1}{c}M^{-1/2}(\mathbf{q}-\mathbf{f})\right] +\frac{1}{2c}(\mathbf{q}(t)-\mathbf{f}(\mathbf{x},\dot{\mathbf{x}}))^TM^{-1}(\mathbf{q}(t)-\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})) -(\dot{\mathbf{x}}+c\mathbf{x})^T\mathbf{g}(\dot{\mathbf{x}}+c\mathbf{x}) \quad (16)$$

$$\text{Let } q_0^2\equiv\text{Sup}[\mathbf{q}(t)-\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})]^TM^{-1}[\mathbf{q}(t)-f(\mathbf{x},\dot{\mathbf{x}})] \quad (17)$$

Then making use of the properties of M, D, K, $\mathbf{g}$,$\mathbf{q}$, and $\mathbf{f}$,

$$\dot{V}\le-\frac{c}{2}[(\dot{\mathbf{x}}+c\mathbf{x})^TM(\dot{\mathbf{x}}+c\mathbf{x}) +2\mathbf{x}^T(K+c^2M)\mathbf{x}+\alpha F(\mathbf{x})]+\frac{q_0^2}{2c} \quad (18)$$

Since $\alpha>1$,

$$\dot{V}\le-cV+\frac{q_0^2}{2c} \quad (19)$$

Let Ω be the bounded set where $V\le q_0^2/2c^2$ and let Ω^c denote the complement of Ω.

Since

$$\dot{V}<0 \text{ for } (\mathbf{x},\dot{\mathbf{x}})\in\Omega^c \quad (20)$$

and since for every finite time interval $0\le t\le a$, $\|\mathbf{x}\|$, $\|\dot{\mathbf{x}}\|\to\infty$, implies $V(\mathbf{x},\dot{\mathbf{x}},t)\to\infty$, the system of equation (12) has no finite escape time. Therefore all solutions are ultimately bounded in Ω and the controlled system is Lagrange stable. Hence, bounded input-bounded output stability is assured if, for the given excitation $\mathbf{q}(t)$, nonlinear function $\mathbf{f}(\mathbf{x},\dot{\mathbf{x}})$, and mass matrix M, a constant $c>0$ can be found such that $V\le q_0^2/(2c^2)$.

4 Applications

The utility of the proposed control strategy is demonstrated by considering two example structures with widely diverse characteristics: (1) a physical model of a three-story building frame which has been extensively analyzed, both analytically and experimentally [10, 11], at the University of California, Berkeley (UCB), and (2) a nonlinear 25-DOF building model that is representative of modern buildings [12].

4.1 Example 1. Subjecting a linearized 3-DOF model of the UCB structure to a representative sample (earthquake D1) of a widely used set of artificial earthquakes [13] and simulating the control strategy discussed in Section 2 yields the results shown in Fig. 3. The three natural frequencies of the model are ≈ 2, 6, and 10 Hz. Thus, the time history segment shown in Fig. 3 corresponds to about 20 fundamental periods T_1 of the system.

The system in Fig. 3 has a sensor and a controller at each of the three mass locations. The pulse duration used here is chosen to equal 0.02 s, approximately equal to 4 percent of T_1. The magnitude of each pulse is proportional to the first power ($n=1$) of the corresponding velocity, with the proportionality constant $c_i=500$, $i=1,2,3$.

Comparison of the peak displacement $y_{i_{\max}}$, the peak velocity $\dot{y}_{i_{\max}}$, the rms value of the displacement $<y_i(t)>$,

$$<y_i(t)>=\{\bar{y}_i^2(t)\}^{1/2}=\left\{\frac{1}{t}\int_0^t y_i^2(\tau)d\tau\right\}^{1/2} \quad (21)$$

and the rms value $<\dot{y}_i(t)>$ of the velocity, with their corresponding values $\hat{y}_{i_{\max}}$, $\hat{\dot{y}}_{i_{\max}}$, $<\hat{y}_i(t)>$, and $<\hat{\dot{y}}_i(t)>$ under active control, shows that the peak rms values of the displacement and velocity of each location are substantially reduced. There is a reduction by a factor of 4 for the peak displacement, a factor of ≈2.5 for the peak velocity, and a factor of ≈6 for the rms displacement and velocity. The results in Fig. 3 and subsequent figures are representative of more than 20 different computer experiments.

The amount of energy needed to accomplish the control strategy can be conveniently gauged by measuring the impulse $I_i=\int p_i(t)\,dt$ of the control force applied to different locations m_i in the structure. In this example, $I_1=1.39\times10^3$, $I_2=1.85\times10^3$, and $I_3=2.08\times10^3$ lb-s, whose ratio of {0.67:0.89:1.0} is very close to the components of the eigenvector associated with the fundamental mode of vibration of the structure: $\varphi^{(1)}=\{0.64, 0.88, 1.0\}$. Clearly, it is more efficient to apply the larger control forces at locations

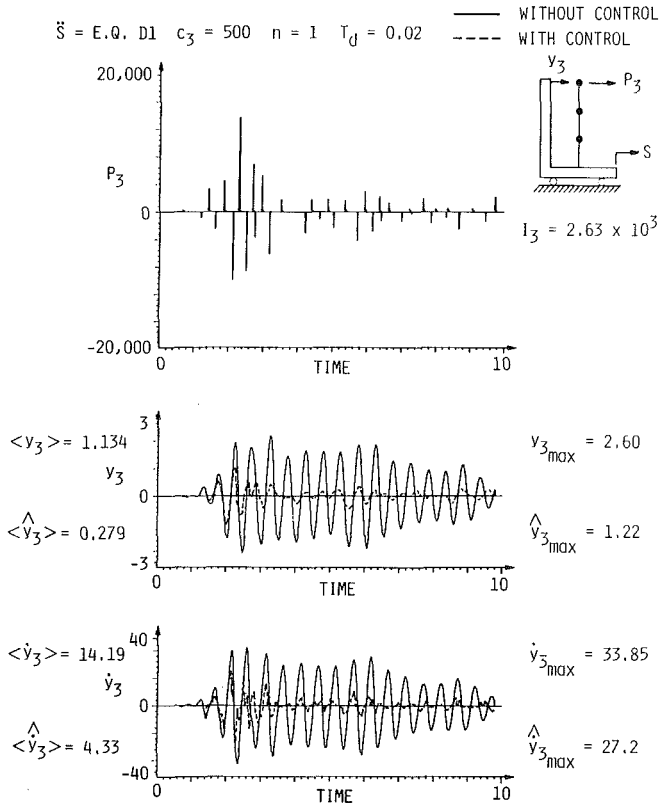


Fig. 4 Response of UCB structure with nonlinear damping; $n=1$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

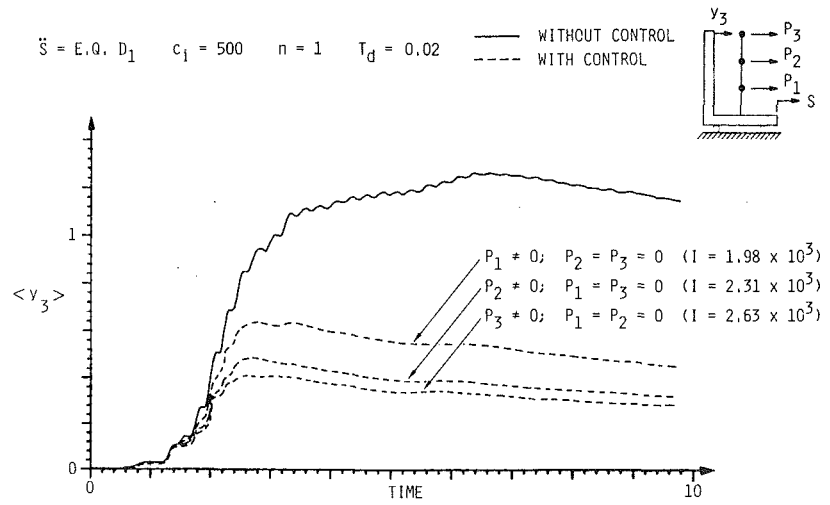


Fig. 5 Effects of controller location on the rms response

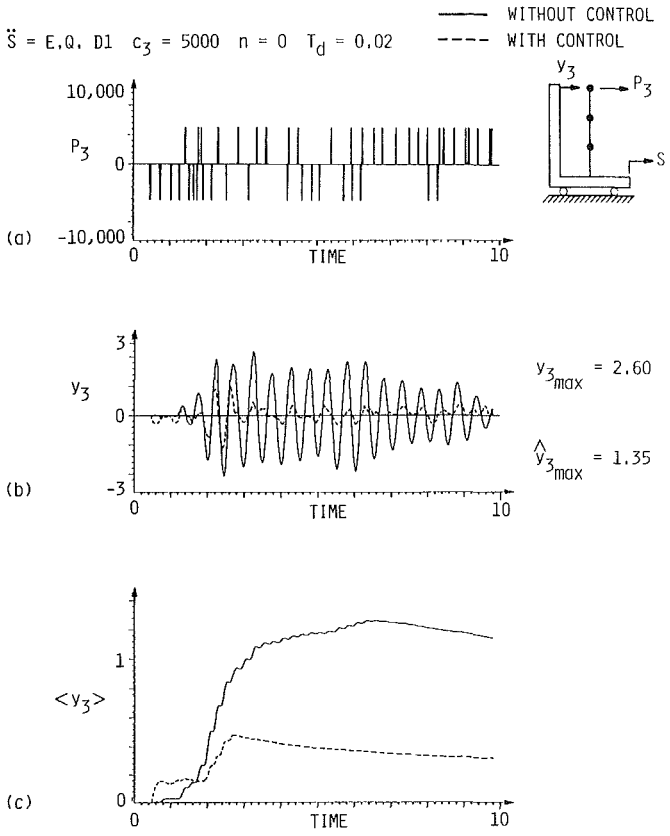


Fig. 6 Reponse of UCB structure with nonlinear damping; $n = 0$

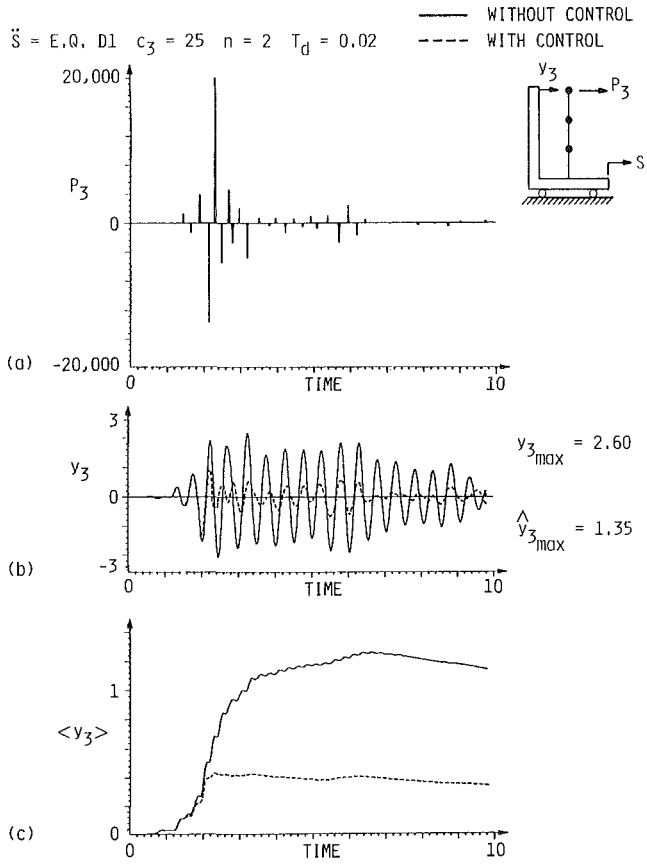


Fig. 7 Response of UCB structure with nonlinear damping; $n = 2$

that can exert more control action by virtue of the amplification effects associated with the effective "moment arm" of each controller location.

The results in Figs. 4 and 5 show what effects can be obtained by using a single controller instead of the three used previously, with all other control and excitation parameters remaining the same. With three controllers, the cumulative impulse used was $\sim 5.3 \times 10^3$, which is reduced to 2.6×10^3, 2.3×10^3, and 2.0×10^3 if a single controller is applied to m_3, m_2, or m_1, respectively. As would be expected, the reduction in the number of controllers, all of similar properties, is accompanied by a lessening in the amount of response reduction. Of course, comparable control results may be achieved by compensating for the reduction in the number of controllers by adjusting the actuator parameters (for example, c_i, n, T_d).

In the cases discussed so far, the actuator force as defined by equation (2) was made proportional to the first power of the corresponding velocity. The results shown in both Figs. 6 and 7 represent a case where a single actuator is applied to the top floor m_3 of the UCB structure. However, in Fig. 6 the pulse amplitude is kept constant (that is, $n = 0$) with $c_3 = 5000$; whereas in Fig. 7 the pulse amplitude is made proportional to the square of the corresponding velocity (that is, $n = 2$) with $c_3 = 25$.

As seen from Fig. 6(a), with $n = 0$, the control force is essentially a Coulomb damping force repetitively introduced

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

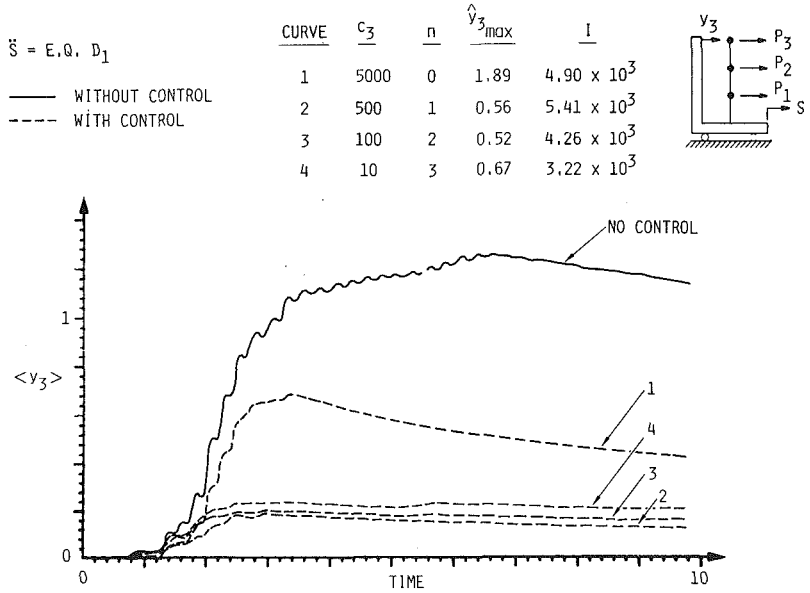


Fig. 8 Comparison of rms response with and without continuous control

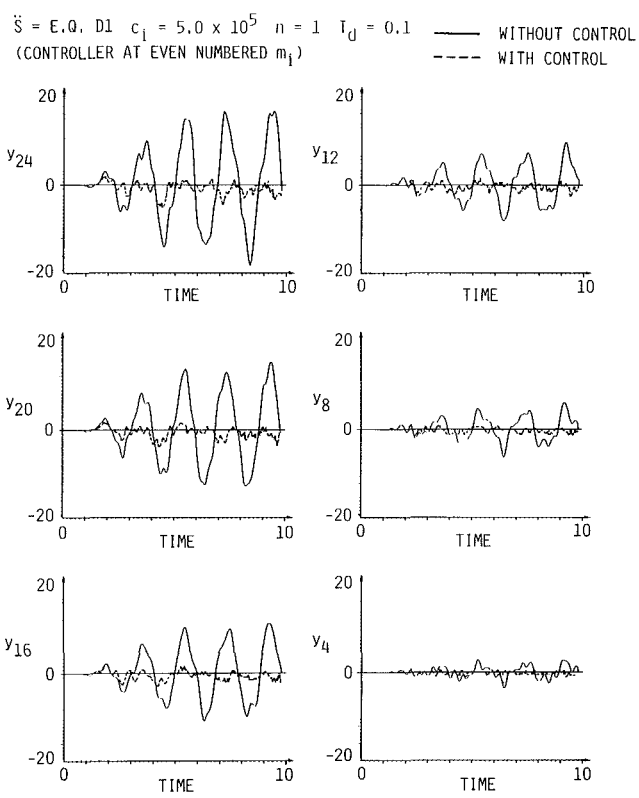


Fig. 9 Controlled response of a linear 25 DOF system

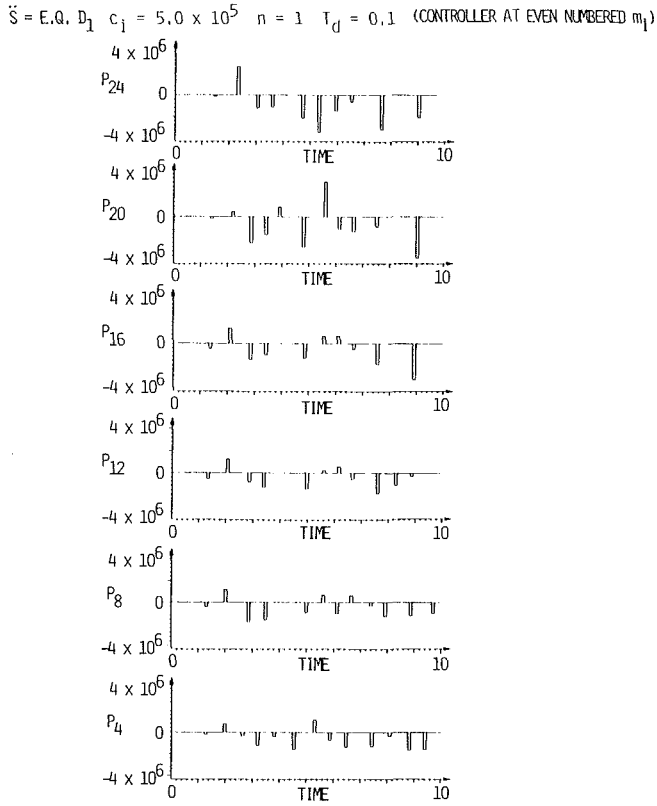


Fig. 10 Control pulses for linear 25 DOF system

at an optimum time (when the velocity is near its peak value), resulting in about 50 percent reduction in the peak response and requiring a total control impulse of $I_3 = 4.60 \times 10^3$lb-s. In the case of Fig. 7, making the control force proportional to a quadratic power of the system velocity results in approximately the same percentage of attenuation of the response with a total control impulse $I_3 = 1.98 \times 10^3$lb-s.

Comparison of the results in Figs. 4, 6, and 7 shows that the adjustment of the coefficient c_i and the velocity power n permits a wide latitude in selecting the actuator control parameters to match hardware limitations. For example, it is generally more efficient, from the point of view of response reduction and control energy utilization, to use as large a value of n as practical. However, the simplicity in hardware design, as well as restrictions on the peak thrust of the actuators may tip the scale toward using a simple device such as the one used in Fig. 6. This actuator, which has a specific thrust capacity, is simply turned on or off at certain times.

To gauge the effects of adjusting the control pulse duration T_d on the response of the structure, several cases were considered in which T_d was set longer than the fundamental period of the structure. Thus, the control strategy was converted from a sequence of discrete control pulses to one with continuous feedback involving nonlinear damping. The

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

results of this investigation are summarized in Fig. 8 for four different representative combinations of c_i and n, representing feedback control with Coulomb, viscous, quadratic, and cubic damping.

Note that in each case there is a significant variation in both the control energy expended and the resulting attenuation of the response. The peak response for steady-state motion with nonlinear velocity-dependent damping is proportional to the inverse of the power of the velocity. It is not surprising, therefore, that under nonstationary random excitation the velocity-dependent control forces are much more efficient than the Coulomb-type in limiting the peak system response while requiring less overall control energy.

4.2 Example 2. The application of the control method to more complicated structures is illustrated by an example linear 25-DOF building [12] which is representative of modern tall buildings.

For this analysis, earthquake D1 was used as a hypothetical disturbance, and identical controllers were assumed at all even-numbered locations. Each controller was set to generate a pulse of duration 0.1s ($\sim 0.05T_1$) and an amplitude of $c_i = 5.0 \times 10^5$ times the first power ($n = 1$) of the velocity at the controller location. Figures 9 and 10 indicate that the reduction is obtained in both the peak and the rms response throughout the structure. Also, the pulse trains applied at the various locations differ in number, phase, and amplitude and more control energy is applied at points that are farthest away from the point of fixity of the structure.

4.3 Example 3. The structures to be controlled in the first two examples were purely linear. To appreciate the versatility of the method, consider a nonlinear version of the building model in [12]. In this case, the interstory restoring force no longer consists of a linear spring k. Instead, each is now made up of a bilinear hysteretic element with the following properties:

x_y = displacement yield level for each element = 1.0 in.

$$\alpha = \frac{\text{stiffness in yielded range}}{\text{stiffness in elastic range}} = 0.5.$$

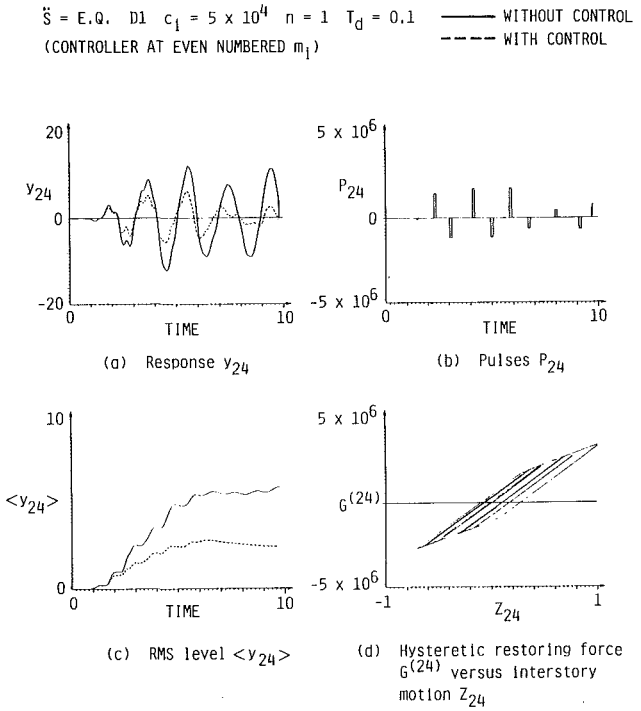


Fig. 11 Controlled response of a nonlinear 25 DOF system

This nonlinear structure is subjected to earthquake D1, assuming the same number and location of controllers as in Example 2. Each of the identical actuators is set to have parameters $c_i = 5 \times 10^4$ and $n = 1$ and to be triggered at the appropriate displacement zero crossing. The resulting response and control forces are shown in Fig. 11. The controlled response of the same nonlinear system with Coulomb-type actuator forces is shown in Fig. 12.

As a further illustration of the flexibility of the proposed control method and its wide applicability to different types of nonlinear distributed systems, the triggering of actuators in Fig. 11 is now tied to the crossing of a preselected velocity threshold. This strategy is useful for dealing with nonlinear structures, such as hysteretic ones or those with time-varying parameters, whose response may include a permanent drift from the structure's initial equilibrium position. Figure 13 shows this procedure to be quite successful in controlling the building response.

5 Discussion

To investigate the dependence of the control procedure on the nature of the excitation, the following test excitations were used: (1) artificial earthquakes A1, B1, and C1 in reference [13]; (2) recorded ground motion corresponding to the 1940 El Centro earthquake [14]; (3) a large ensemble of stationary random excitations; and (4) swept-sine excitation. These excitations, which varied widely in amplitude, duration, and nature, were applied by simulation techniques to the three example structures and to several others of different complexity. Additionally, experimental studies were conducted by an analog computer to simulate the control method applied to different structures under a variety of dynamic environments. In all these cases, the control method consistently attenuated the response of the tested system. Provided that the proper phase relationship between the control pulses and the velocity is maintained, all the systems tested remained stable for all variations of physical parameters, including pulse amplitude and duration. For the sake of brevity, results from only one sample excitation are presented in this paper. Further details are available in [15].

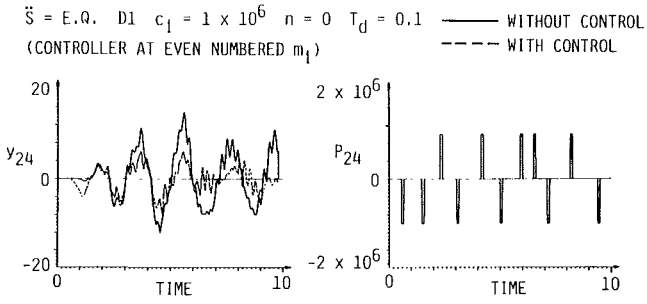


Fig. 12 Controlled response of a nonlinear 25 DOF system

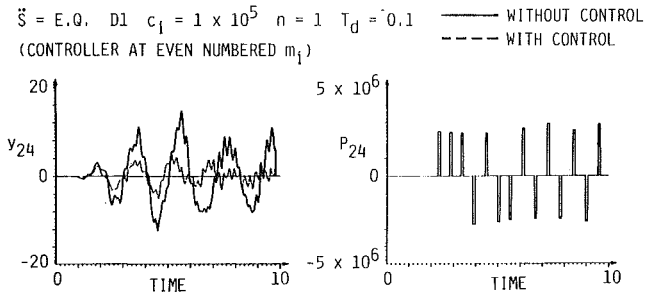


Fig. 13 Controlled response of a nonlinear 25 DOF system using velocity threshold

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

It should be pointed out that while no algorithm is given for the choice of the control pulse parameters, c_i, n_i, and T_{d_i}, their selection is governed primarily by hardware limitations. Thus the available actuator power will dictate the practical limit of the exponent n_i. While tradeoffs are possible between pulse duration and amplitude, nevertheless hardware limitations again impose bounds such as minimum T_{d_i} (related to minimum equipment response time). Clearly, total available control energy will limit the upper bound on both pulse amplitude (c_i) and duration T_{d_i}.

While further work is needed to establish the optimum values of these control parameters, this investigation has clearly demonstrated the feasibility and robustness of the method.

6 Summary and Conclusions

An analytical and "experimental" study was made of the characteristics of a very simple active-control method. The control method is based on the use of pulse actuators attached to selected positions in the structure, whose activation time and amplitude are determined by the local structural response in the vicinity of the controllers.

A stability analysis by means of Liapunov's direct method of a nonlinear n-degree-of-freedom system shows that conditions exist under which the response of the controlled structure for arbitrary excitations is bounded. Simulation studies were performed on linear and nonlinear example structures representative of modern tall buildings under the action of nonstationary random excitations resembling earthquake ground motion. The effects of various control parameters such as actuator locations, numbers, force levels, and triggering strategy were investigated.

The results of these studies show the method to be simple, efficient, and suitable for on-line implementation with realistic nonlinear flexible structural systems subjected to arbitrary dynamic environments.

Acknowledgments

This study was supported in part by a grant from the National Science Foundation. The assistance of T.J. Dehghanyar in the simulation studies is appreciated.

References

1 Robinson, A. C., "A Survey of Optimal Control of Distributed-Parameter Systems," *Automatica*, Vol. 7, 1971, pp. 371–388.

2 Nordell, W. J., "Active Systems for Blast-Resisting Structures," Technical Report R-611, Naval Civil Engineering Laboratory, Port Hueneme, Calif., Feb. 1969.

3 Yao, J. T. P., "Concept of Structural Control, "*Journal of the Structural Division*, ASCE, Vol. 98, No. ST7, July 1972, pp. 1569–1574.

4 Yang, J. N., "Application of Optimal Control Theory to Civil Engineering Structures," *Journal of Engineering Mechanics Division*, ASCE, Vol. 101, No. EM6, Dec. 1975, pp. 819–838.

5 Martin, C., and Soong, T. T., "Modal Control of Multistory Structures," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 102, No. EM4, Aug. 1976, pp. 613–623.

6 Abdel-Rohman, M., and Leipholz, H. H. E., "Structural Control by Pole Assignment Method," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 104, No. EM5, Oct. 1978, pp. 1159–1175.

7 Leipholz, H. H. E., ed., *Structural Control*, North Holland SM Publications, 1980.

8 Masri, S. F., Bekey, G. A., and Caughey, T. K., "Optimum Pulse Control of Flexible Structures," ASME, JOURNAL OF APPLIED MECHANICS, Vol. 48, Sept. 1981, pp. 619–626.

9 National Science Foundation (NSF), *Feasibility of Force Pulse Generators for Earthquake Simulations*, NSF Grant PFR 77-15010, Washington, D.C., NSF, 1979.

10 Clough, R. W., and Tang, D. T., "Earthquake Simulated Study of a Steel Frame Structure, Vol. I: Experimental Results," Report No. EERC 75-6, University of California, Berkeley, Calif., Apr. 1975.

11 Tang, D. T., "Earthquake Simulated Study of a Steel Frame Structure, Vol. II: Analytical Results," Report No. EERC 75-36, University of California, Berkeley, Calif., Oct. 1975.

12 Blume, J. A., Newmark, N. M., and Corning, L. H., *Design of Multistory Reinforced Concrete Buildings for Earthquake Motions,* Portland Cement Association, Skokie, Ill., 1961.

13 Jennings, P. C., Housner, G. W., and Tsai, N. C., "Simulated Earthquake Motions," *California Institute of Technology Report*, 1968.

14 Hudson, D. E., "Destructive Earthquake Ground Motions," in: *Applied Mechanics in Earthquake Engineering*, ASME, New York, 1974.

15 Masri, S. F., Bekey, G. A., and Caughey, T. K., "On-Line Control of Nonlinear Flexible Structures," University of Southern California Report USC-CE-8106, Apr. 1981.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

C. S. Hsu
Professor of Applied Mechanics,
Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720
Fellow ASME

R. S. Guttalu
Research Associate,
Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720

W. H. Zhu
Department of Mechanics,
Zhejiang University,
Hangzhou, Zhejiang,
People's Republic of China

A Method of Analyzing Generalized Cell Mappings

In this paper a method for analyzing the global properties of generalized cell mappings involving a very large number of cells is presented. The procedure of the algorithm is devised by taking advantages of certain properties of Markov chains. First, the persistent groups representing the motions to which the system will eventually evolve are determined. Then the absorption probabilities of any transient cell into these persistent groups are evaluated. These two sets of information provide us with the global behavior of the system. After describing the method its viability is illustrated by two examples.

1 Introduction

Based on the concept of a cellular state space, a theory of cell-to-cell mappings has been developed in [1-3] for the global analysis of nonlinear systems. In [1, 2] a primitive theory of cell-to-cell mappings is offered. A cell $\mathbf{Z}(n)$ at $t=n$ is mapped by a cell mapping $\mathbf{C}$ into cell $\mathbf{Z}(n+1)$ at $t=n+1$.

$$\mathbf{Z}(n+1)=\mathbf{C}(\mathbf{Z}(n)) \tag{1}$$

Using this approach, the determination of the domains of attraction becomes a matter of sorting the cells and it leads to an extremely efficient method of analysis. This primitive cell-to-cell mapping theory does have its drawback in its crudeness unless very small cells are used. Mainly to improve on this, a generalized cell mapping theory is presented in [3]. In the generalized theory a cell under the mapping is allowed to have several image cells with each image cell having a definite fraction of the total probability. This formulation leads directly to Markov chains in the form

$$\zeta(n+1)=\mathbf{P}\zeta(n) \tag{2}$$

where $\zeta(n)$ is the cell probability vector at $t=n$ and $\mathbf{P}$ is the transition probability matrix. Some simple examples involving only very small number of cells were offered in [3] but applications of the theory using a very large number of cells awaited the development of an appropriate algorithm. In this paper we offer such an algorithm. With this algorithm at hand we believe we have a new approach to the global analysis of nonlinear systems.

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February, 1982. Paper No. 82-WA/APM-23.

Copies will be available until July, 1983.

After describing the algorithm in Sections 2-9, some examples of application are given in Section 10 to demonstrate the feasibility of the method. Throughout the paper we shall refer constantly to [1-3], particularly [3]. For the terminology used in this paper for Markov chains the reader is referred to [4-6].

2 Outline of the General Procedure

The general procedure to be adopted here is very similar to the algorithm described in [2], except that now, in addition to sorting, which is really the only operation involved in the algorithm of [2], we must also evaluate the propagation of the probability distribution among the cells as the system evolves. The description of the algorithm will be given in two parts. In part one we present the basic algorithm where the emphasis will be on describing the basic ideas without being overly concerned about the aspect of efficiency. In part two we describe various possible steps one can take to improve the efficiency of the algorithm.

The basic algorithm again may be divided into several parts. In the first part we process the cells in order to classify them as *persistent* or *transient* cells. For the persistent cells we also decompose them into *irreducible persistent groups*. In the second part we determine the period of each persistent group and also the long-term *limiting probability distribution* among the cells within each group by using Theorem 4 of [3] for the *acyclic groups* and Theorem 7 of [3] for the *periodic groups*. In the third part we evaluate the evolution from all the transient cells by computing the *absorption probability* α_{ij} of a transient cell j into a persistent cell i, the *groups absorption probability* $\alpha(\mathbf{B}_g, j)$ of cell j into a persistent group $\mathbf{B}_g$, and the *expected absorption time* ν_j of cell j into persistent groups. For this purpose we first decompose the set of transient cells into transient groups so that the normal form of the transition probability matrix is in the form (17) of [3].

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The cell mapping will be specified in the following manner. Let Z be the cell designation number. For the purpose of our present discussion we designate the sink cell as cell 0 and the regular cells as cells 1, 2,..., N_c. In actual computer implementation one may need to designate the sink cell as cell (N_c+1) because the label 0 may not be available. In general, a cell Z may be mapped into several cells. Let $I(Z)$ denote the number of image cells of Z. Next, we use two arrays of mapping data $C(Z, i)$ and $P(Z, i)$ where Z ranges from 0 to N_c while i from 1 to $I(Z)$. $C(Z, i)$ designates the ith image cell of Z and $P(Z, i)$ is the transition probability from cell Z to its ith image cell in one mapping step. Obviously we should always have

$$\sum_{i=1}^{I(Z)} P(Z,i)=1 \quad \text{for all} \quad Z. \tag{3}$$

These three arrays $I(Z)$, $C(Z, i)$, and $P(Z, i)$ *constitute* the input data to this global analysis algorithm for generalized cell mappings.

Each persistent group will be given a number designation called the *group number*. The groups will be numbered sequentially, 1, 2,..., G_p, as they are discovered. To each persistent cell is assigned the group number of the persistent group to which it belongs and the notation is $G(Z)$. A persistent group with a group number g will be called the gth persistent group and, when considered as a set of cells, will be denoted by $\mathbf{B}_g$. The total number of member cells in this gth group will be denoted by the *member number* $N(g)$. Each persistent group has, of course, a period. We use $K(g)$, the *period number*, to denote the periodicity of the gth persistent group[1]. An absorbing cell and an acyclic group will have their periodicity numbers equal to 1. To a large extent the numbers $N(g)$ and $K(g)$ characterize a persistent group.

Each transient group and all its members will also be given a group number. These groups will also be numbered sequentially as they are discovered. Since in this algorithm the transient groups are always determined after all the persistent groups have been discovered, their group numbers follow after G_p; namely: G_p+1, G_p+2, ..., G_p+G_t, where G_t is the total number of the transient groups. Thus, the hth transient group will have a group number equal to G_p+h. To each transient group we also assign a *member number* $N(G_p+h)$ to denote the number of the members in the group. With regard to the periodicity number, a transient group does not have a period. However, for the convenience of identification we *arbitrarily* assign -1 as the periodicity number of any transient group, i.e., $K(G_p+h)=-1$ for $h=1, 2,..,G_t$. With these notation conventions we can read the classification of a cell in the following manner once the classification work has been completed. For a cell Z we find $G(Z)$. If $G(Z) \leqq G_p$, Z is a persistent cell and if $G(Z)>G_p$, Z is a transient cell. Or, one can find $K(G(Z))$. If it is a positive integer d then Z is a persistent cell of period d. If it is equal to -1, Z is a transient cell. A redundancy of information is provided here.

When the persistent and transient groups are taken together their member numbers must satisfy

$$\sum_{i=1}^{G_p+G_t} N(i)=N_c+1 \tag{4}$$

If a transient cell Z leads to a persistent group $\mathbf{B}_g$, then we call $\mathbf{B}_g$ a *domicile* of cell Z. A transient cell may have several domiciles. We denote the total number of the domiciles of a cell Z by its *domicile number* $D(Z)$. Of course $D(Z)$ cannot exceed G_p. To identify the corresponding persistent groups as the domiciles we use the notation $DG(Z, i)$, $i=1,2,..,D(Z)$, to denote the persistent group which is the ith domicile of cell Z. For the sake of uniformity we also extend the preceding definition to cells that are persistent. For a persistent cell Z belonging to a persistent group $\mathbf{B}_g$, it is natural for us to set $D(Z)=1$ and $DG(Z, 1)=g$.

In [3] the probability of a transient cell Z to be absorbed into a persistent Z' is denoted by[2] $\alpha_{Z'Z}$. Let $\mathbf{B}_g$ be the jth domicile of Z and let Z' be a member of $\mathbf{B}_g$. Then $\alpha(\mathbf{B}_g, Z)$ defined by (57) in [3] is the group absorption probability from a transient cell Z into its jth domicile. Not all persistent groups are the domiciles of a transient cell. For an efficient algorithm it is better to ignore, for each Z, the persistent groups that are not its domiciles. For this reason we use a new notation $A_1(Z, j)$ to denote the *group absorption probability* of a *transient cell* Z into its jth *domicile*.

$$A_1(Z,j)=\sum_{Z'\in\mathbf{B}_{DG(Z,j)}} \alpha_{Z'Z} \tag{5}$$

The index j in $A_1(Z, j)$ ranges from 1–$D(Z)$. For a persistent cell its domicile number is 1 and, strictly speaking, $A_1(Z, 1)$ for such a cell is not needed. In the algorithm we arbitrarily set $A_1(Z, 1)=1$ for a persistent cell Z.

Next, consider the expected absorption time. Let Z be a transient cell and Z' be a member of its jth domicile. Then one can evaluate the expected absorption time of Z into Z' under the condition that Z is absorbed into Z'. This "conditional expected (or mean) absorption time" of Z into Z' will be denoted by $\nu_{Z'Z}$.

$$\nu_{Z'Z}=\sum_{n=1}^{\infty} n f_{Z'Z}^{(n)}/\alpha_{Z'Z} \tag{6}$$

The expected absorption time of Z into its jth domicile $\mathbf{B}_g$ under the condition that Z is absorbed into $\mathbf{B}_g$ will be called "conditional group expected absorption time" of Z into its jth domicile and is denoted by $\nu(Z, j)$.

$$\nu(Z,j)=\sum_{Z'\in\mathbf{B}_g}\sum_{n=1}^{\infty}\frac{n f_{Z'Z}^{(n)}}{A_1(Z,j)}$$

$$=\sum_{Z'\in\mathbf{B}_g}\nu_{Z'Z}\frac{\alpha_{Z'Z}}{A_1(Z,j)} \tag{7}$$

The *expected absorption time* of a transient cell Z into all its domiciles is given by

$$\nu_Z=\sum_{j=1}^{D(Z)}\sum_{Z'\in\mathbf{B}_g}\sum_{n=1}^{\infty} n f_{Z'Z}^{(n)}$$

$$=\sum_{j=1}^{D(Z)}\sum_{Z'\in\mathbf{B}_g}\nu_{Z'Z}\alpha_{Z'Z}$$

$$=\sum_{j=1}^{D(Z)}\nu(Z,j)A_1(Z,j), \tag{8}$$

where $\mathbf{B}_g$ is understood to be the jth domicile of Z. Again, for a persistent cell Z, $\nu(Z, j)$ and ν_Z are not defined; however, in the algorithm we arbitrarily set $\nu(Z, 1)=0$ and $\nu_Z=0$.

For the global analysis, after having determined the persistent groups, the data $D(Z)$, $A_1(Z, j)$, and ν_Z for the transient cells are the key results. They state that a transient

[1] With regard to $K(g)$, the reader may find that in [2] a periodicity number, denoted by $P(Z)$ in [2], is assigned to each cell. This is convenient and also makes sense there because in the theory of simple cell mapping each cell belongs to only the domain of attraction of one periodic solution. For the present generalized cell mappings, a transient cell may belong to the domains of attraction of several persistent groups; therefore, no single periodicity number can be assigned to a transient cell in general.

[2] In value $\alpha_{Z'Z}=f^*_{Z'Z}$. In [3] a different symbol $\alpha_{Z'Z}$ has been introduced to emphasize the transition from a transient cell to a persistent cell.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

cell Z has $D(Z)$ number of domiciles, it is absorbed into its jth domicile with a probability $A_1(Z, j)$, and on the average is absorbed into the persistent groups in ν_Z steps of time.

3 Absorbing Cells

An easy task, which should be done first, is to locate all the absorbing cells. A cell Z is an absorbing cell if and only if

$$I(Z)=1 \quad \text{and} \quad C(Z,1)=Z. \tag{9}$$

Of course, in this case $P(Z, 1)=1$ if the input data is not in error. Thus, we can process the cells sequentially one-by-one to find the absorbing cells by testing (9). To characterize the absorbing cells once they are found we make use of the group number $G(Z)$, the member number $N(g)$, the periodicity number $K(g)$, the domicile number $D(Z)$, the group absorption probability $A_1(Z, j)$, and the expected absorption time ν_Z. For example, we begin the process by examining $Z=0$. As $I(0)=1$ and $C(0, 1)=0$ because cell 0 is the sink cell, $Z=0$ is an absorbing cell. Thus, we have found the first persistent group in the form of an absorbing cell. Cell 0 is then assigned with $G(0)=1$, $N(1)=1$, $K(1)=1$, $D(0)=1$, $A_1(0, 1)=1$, and $\nu_0=0$. Next, we successively examine $Z=1, 2,\ldots,N_c$ to see whether the conditions of (9) are satisfied. Each time an absorbing cell is found, a new persistent group is discovered and the next higher group number is assigned to that cell and the various numbers are assigned accordingly. In this one sweep we locate all the absorbing cells.

4 Search for Other Persistent Groups

Next, we search for acyclic and periodic groups of persistent cells and, in the process, also determine all the transient cells. The algorithm is patterned after that given in [2]. It is based on some key properties of finite Markov chains.

(*i*) A cell is either a persistent cell belonging to a persistent group, or it is a transient cell.

(*ii*) Starting from a persistent cell of a particular persistent group, the system can only go to fellow-member cells of that group and cannot go to a transient cell or to a persistent cell of another group.

(*iii*) If a cell leads to a persistent cell of a particular persistent group, say $\mathbf{B}_g$, and yet is not a member of that persistent group, then it is a transient cell and has $\mathbf{B}_g$ as one of its domiciles.

(*iv*) If a cell Z leads to cell Z', which is known to be transient, then Z is transient and has all the domiciles of Z' as its own domiciles. Z can, of course, have additional domiciles.

Again, as in [2], the algorithm involves repeatedly calling a new cell from the cell space, following its mapping evolution to generate a cell sequence, and processing the sequence in such a way that all the cells in the sequence can be classified. There are three possibilities for any processing sequence. It may lead to a cell that is *known* to be a transient cell, then according to (*iv*) all the cells in the sequence are transient cells and they can be so tagged. The sequence may lead to a cell that is *known* to be a persistent cell of a previously discovered persistent group. In that case, according to (*iii*), again all the cells in the sequence are transient and are to be so tagged. The third possibility is that it leads to a new persistent cell. Once such a new persistent cell has been discovered, we can continue the sequence in a special way to locate all the cells to which it leads. This, according to (*ii*), will give us all the members of this new persistent group. All the other cells in the sequence are necessarily, according to (*iii*), transient. These persistent and transient cells can then be appropriately tagged. Once the classification of the cells has been completed for a sequence, we begin[3] a new sequence in order to process the other cells.

In the processing of the said sequence we need to distinguish four kinds of cells. First we have the *virgin cells*. They are the cells that have never been called before in the current or previous sequence. As the group number $G(Z)$ of all cells are set to zero at the beginning of a run of the program, a virgin cell Z is characterized by having $G(Z)=0$. Cells of the second kind are those that have been called in the current sequence. These will be called *cells under processing*. If we adopt the rule that a virgin cell once called will have its group number changed to -1, then all the cells under processing will be characterized by group number -1. At a certain stage of generating the sequence, we need to test to see whether a cell Z is indeed a persistent cell. Such a cell will be called a *persistent cell candidate* or simply a *candidate cell*. When a cell under processing becomes a candidate cell, its group number is temporarily changed from -1 to -3 for the identification purpose. Of course, after testing, a persistent cell candidate may turn out not to be one; in that case its group number is reverted back to -1 and the sequence proceeds forward. Finally, once a new group has been discovered and determined, then a group number, say g, is assigned to that group and all member cells of the group will have their group number changed permanently from -3 to g. All the other cells in this sequence are transient cells and will have their group numbers changed semipermanently from -1 to -2. Here -2 for the group number is simply used as an identification symbol for a transient cell until the transient cell itself is classified into a transient group at a later stage of the program. These cells with permanent or semipermanent group numbers are the fourth kind of cells and they are called the *processed cells*. Thus, group numbers, permanent or temporary, have the following meanings:

$G(Z)=-3$, Z is a persistent cell candidate, a temporary designation;
$G(Z)=-2$, Z is a transient cell, a semipermanent designation;
$G(Z)=-1$, Z is a cell under processing, a temporary designation;
$G(Z)=0$, Z is a virgin cell, a temporary designation;
$G(Z)=g\geq 1$, Z is a member of the gth persistent group, a permanent designation. (10)

We now describe in detail how a processing sequence is generated. We start with a virgin cell Z. We set $G(Z)=-1$ to indicate that it is now in a current processing sequence. For notation convenience we designate Z as $B(0)$. Next we find the *first* image of $B(0)$. If $C(B(0),1)$ is a virgin cell, we call it $B(1)$, set $G(B(1))=-1$, and proceed to the first image cell of $B(1)$. Continuing in this manner we generated the following sequence:

$$\begin{array}{ccccccc} Z & \rightarrow C(B(0),1) & \rightarrow C(B(1),1) & \rightarrow \ldots & \rightarrow C(B(M_1-1),1) & \rightarrow C(B(M_1),1) \\ \| & \| & \| & & \| & \\ B(0) & \rightarrow B(1) & \rightarrow B(2) & \rightarrow \ldots & \rightarrow B(M_1) & \end{array} \tag{11}$$

The total number of cells being finite, sooner or later in this sequence a cell will appear that is either a *processed cell* or a

[3] In the actual algorithm, when a new persistent group is discovered we immediately process the group in order to determine its period and its limiting probability distribution before proceeding to start a new sequence to locate the next persistent group. This work of determination of the properties of the groups will be discussed in the next two sections.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

cell under processing. Let this cell be $C(B(M_1),1)$ in the sequence.

(*A*) If $C(B(M_1),1)$ is a processed cell, then no matter whether it is persistent cell (with a posititve integer as its group number) or a transient cell (with -2 as its group number), all the preceding cells in the current sequence are, according to (*iii*) and (*iv*), transient cells. In this case we set the group numbers of $B(0)$, $B(1),\ldots,B(M_1)$ all equal to -2. These cells have now been classified and they have become processed cells. The current sequence is terminated and we go back to the cell space to pick the next virgin cell to start a new processing sequence.

(*B*) The second possibility is that $C(B(M_1),1)$ is a cell under processing, i.e., it has already appeared as one of the preceding cells in the sequence. In this case a continuation of the sequence along the "*first image track*" will simply duplicate the cycle indefinitely and we say a "*looping*" cell has been encountered. Now a sequence modification is necessary. Here we temporarily halt the sequence and go back to check whether $B(M_1)$ has one or more image cells.

(*B-I*). If $B(M_1)$ has more than one image cells, i.e., $I(B(M_1))>1$, then we continue but modify the sequence by replacing the cell $C(B(M_1),1)$ at the nominal (M_1+1)th position in the sequence by $C(B(M_1),2)$, the second image cell of $B(M_1)$.

$$\begin{array}{l}\ldots\rightarrow C(B(M_1-1),1)\rightarrow C(B(M_1),2)\\ \qquad\quad\|\\ \ldots\rightarrow B(M_1)\end{array} \tag{12}$$

Again, one has three possibilities:

(*B-I-1*). $C(B(M_1),2)$ is a processed cell, then the situation (*A*) applies. All the cells in the current sequence are transient cells. After tagging, the sequence is terminated.

(*B-I-2*). $C(B(M_1),2)$ is a virgin cell. In that case we call it $B(M_1+1)$, set $G(B(M_1+1))=-1$, and restart the sequence by finding $C(B(M_1+1),1)$ as the next member, i.e., restarting the sequence by reverting back to the *first image track*.

$$\begin{array}{llll}\ldots\rightarrow C(B(M_1-1),1) & \rightarrow C(B(M_1),2)\rightarrow C(B(M_1+1),1)\\ \quad\|\\ \ldots\rightarrow B(M_1) & \rightarrow B(M_1+1)\end{array} \tag{13}$$

The sequence is then continued forward as in (11).

(*B-I-3*). $C(B(M_1),2)$ is a cell under processing. In that case we again encounter a looping cell and we modify the sequence by replacing the looping image cell (here the second) by the next higher image cell (here the third) if it exists. The checking of the three possibilities for the new image cell is then to be repeated.

(*B-II*). Suppose that at one stage of the sequence development one finds that all the image cells of a cell, say $B(M)$, are cells under processing. Then potentially we have the discovery of a persistent group. Here M may be M_1, the position in the sequence where the first looping cell is encountered, or it could be a later position in the sequence where another looping cell is encountered.

Testing for a Possible Persistent Group. A persistent group is characterized by being a closed set of communicating cells. This means all the image cells of every cell in the set belong to the set. When $B(M)$ and all of its image cells are cells under processing, then these cells might be the members of a persistent group. However, this possibility needs to be tested and confirmed. To do so, we first identify these persistent cell candidates by changing their group numbers temporarily to -3. We refer to this as the *candidate set*. At this time the set may not be complete and it may need to be supplemented with additional candidate cells at later stages, or it may turn out to be a false set. The testing involves systematically picking a member from this candidate set and seeing whether all its images cells are in the set. Several possibilities can happen. (1) The tested image cell of a candidate cell is a candidate cell. In that case we continue the testing forward. (2) The tested image cell is a cell under processing but not in the current candidate set. Such a cell is also a possible persistent cell candidate. Therefore, we enlarge the candidate set by adding this cell to the set, and continue the testing. (3) The tested cell is a processed cell. In that case the test has failed and all cells in the processing sequence are transient cells. The cells will be so tagged and the sequence is terminated. (4) The tested cell turns out to be a virgin cell. In that case we get out the testing subroutine and continue the main processing sequence by restarting the sequence with this virgin cell as in (11). Before doing so, we first, of course, need to change the tagging of all the candidate cells back to cells under processing.

In this manner we can locate the persistent group if one is indeed contained in the sequence. When a persistent group is discovered, it is assigned a new group number, say the gth group, and this number is then given to every member of the group as its group number. At the same time the member number $N(g)$ of this group is assigned. All the other cells in the processing sequence are transient and will be tagged with -2 as their group numbers.

When the cell space **S** has been exhausted by the processing sequences, then we would have classified all the cells. Each cell is either a persistent cell belonging to a particular persistent group or a transient cell.

5 Determination of the Period of a Persistent Group

After finding a persistent group, the next task is to determine its period. One way to do this is to use Theorem 10 of [3]. If a persistent group is of period d its transition probability matrix must have dth roots of unity among its eigenvalues. To solve an eigenvalue problem for a large matrix is however a time-consuming task. Another way is to test systematically whether $\exp(2\pi i/k)$ is an eigenvalue of the transition probability matrix for $k=N(g), N(g)-1,\ldots,2$, where g denotes the group number of the group. However, this checking involving computation of determinants is also time consuming if $N(g)$ is large. We propose here a procedure that is based on the result that a persistent group of period d has d disjointed *subgroups* and permits a normal form of the transition probability matrix in the form of (25) of [3].

Let the group under consideration be the gth group and let its members be labeled as $Y_1, Y_2,\ldots,Y_{N(g)}$. This will be referred to as the Y-set. We also adopt the notation $M(g,i,j)$ as the jth member of the ith subgroup of the gth persistent group, and $N(g,i)$ as the number of members in the ith subgroup of this gth persistent group. One has

$$\sum_{i=1}^{d} N(g,i)=N(g) \tag{14}$$

According to the definition of the term period, if any cell of the group maps into itself, then the period is immediately one and we have an acyclic group. This simple feature is to be implemented at the beginning of the subroutine.

The remainder of the subroutine program is so devised as to take advantage of the property that for a persistent group of period d each cell can only belong to one of the d subgroups. Thus, all the image cells of a cell, say Y_i, belong to one subgroup. Moreover, if another cell Y_j has an image cell belonging to this subgroup, then all the image cells of Y_j belong to this same subgroup. It also follows that Y_i and Y_j belong to one subgroup. After the membership of one subgroup has been determined, the other subgroup can be generated by successive mappings. We now describe the basic steps of the subroutine. In it we use $SG(Y_i)$ to denote the subgroup to which Y_i belongs. $SG(Y_i)=0$ means the

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

subgroup number has not yet been assigned and Y_i is called an unassigned cell.

(1) We start with Y_1 of the Y-set and set

$$SG(C(Y_1,i))=1, \quad i=1,2...,I(Y_1). \tag{15}$$

By (15) we assign all image cells of Y_1 to be in the first subgroup. In general, this will not give us the complete membership of this subgroup. Our first task is to complete the membership of this first subgroup.

(2) Next we take Y_2 from the Y-set and check its image cell set.

(3-*i*) If all the image cells of Y_2 are unassigned cells, we set aside Y_2. We go on to the next cell Y_3 of the Y-set and repeat Step (2). However, before doing so we set $BP(1)=Y_2$ to indicate that Y_2 has been bypassed for the time being in checking the members of the first subgroup. $BP(1)$ means this is the first bypassed cell.

(3-*ii*) If one of the image cells of Y_2 is a member of the first subgroup, indicated by having its SG-number equal to 1, then all image cells of Y_2 are in that subgroup. We add the new members to the subgroup by assigning

$$SG(C(Y_2,i))=1, \quad i=1,2,...,I(Y_2). \tag{16}$$

After having done this we proceed to the next cell in the Y-set and repeat Step (2).

(4) Steps (2) and (3) are to be repeated for all cells in the Y-set. At the end of the sweep we would have discovered, say N_1 members of the first subgroup and left behind, say, B_1 bypassed cells.

(5) Next, we sweep through the bypassed cells in order to discover additional members of the subgroup. If there are additions, then the set of bypassed cells is altered and updated.

(6) The sweeping through the set of bypassed cells is to be repeated until during one sweep there is no change in the membership of the first subgroup. The complete subgroup has now been found.

(7) Let $N(g,1)$ be the total number of cells in this first subgroup. It may happen that $N(g,1) = N(g)$, then this gth persistent group is an acyclic group and our task is completed. If $N(g,1) \neq N(g)$, we tag the member cells of the first subgroup by $M(g,1,j), j=1,2,...,N(g,1)$.

(8) Having found the first subgroup, the second subgroup is easily found. It is simply the set of all the nonrepeating image cells of all members of the first subgroup. They are

$$C(M(g,1,j),h), \quad j=1,2,...,N(g,1), \quad h=1,2,...,I(M(g,1,j)). \tag{17}$$

The number of members of this subgroup is denoted by $N(g,2)$ and the members are tagged by $M(g,2,j)$, $j=1,2,...,N(g,2)$. If it happens that $N(g,1) + N(g,2) = N(g)$, then the persistent group has a period 2 and the task is completed.

(9) If $N(g,1) + N(g,2) \neq N(g)$, then we proceed to locate the third subgroup and so forth.

In this manner the period d, and d subgroups, and their respective memberships can be determined. To complete the description of this persistent group we set $K(g)=d$ to show that the period of the gth group is d.

6 The Limiting Probability Distribution of a Persistent Group

Having determined the persistent groups and their periods, the next task is to find the limiting probability distributions. For an acyclic group the limiting probability distribution is simply the right-eigenvector $\mathbf{p}$ of the transition probability matrix[4] $\mathbf{P}$ associated with eigenvalue 1, i.e.,

$$\mathbf{Pp}=\mathbf{p}, \quad \text{or} \quad (\mathbf{P}-\mathbf{I})\mathbf{p}=\mathbf{0} \tag{18}$$

subjected to the normalization condition

$$\sum_{i=1}^{N} p_i = 1 \tag{19}$$

Let us rewrite (19) in the form

$$p_N = 1 - \sum_{i=1}^{N-1} p_i \tag{20}$$

and the first $N-1$ equations of (18) as

$$\sum_{j=1}^{N-1} (p_{ij} - \delta_{ij} - p_{iN}) p_j = -p_{iN}, \quad i=1,2,...,N-1. \tag{21}$$

This is a set of linear equations in $N-1$ unknowns p_j, $j=1,2,...,N-1$. This can be solved readily using any efficient computer subroutine, including some iteration procedures. Once these components of $\mathbf{p}$ have been found, p_N is found from (20).

Next consider a periodic persistent group of period $d>1$. Let the d subgroups be denoted by $\mathbf{B}_1, \mathbf{B}_2,...,\mathbf{B}_h,...,\mathbf{B}_d$. Let the numbers of cells in the subgroups be denoted by $N_1, N_2,...,N_h,...,N_d$. For the following discussion in this section we assume that the cells are so ordered and designated that $\mathbf{P}$ is in its normal form (25) of [3]. Moreover, by the discussion given in [3] just ahead of Theorem 7, we know that while each $\mathbf{B}_h$ is a subgroup of a periodic persistent group under the mapping action $\mathbf{P}$, it becomes an isolated acyclic group under the action of mapping $\mathbf{R}=\mathbf{P}^d$. $\mathbf{R}$ admits a normal form of

$$\mathbf{R} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{P}_d \end{bmatrix} \tag{22}$$

which has only square block matrices in its diagonal positions and where $\mathbf{P}_h$, $h=1,2,...,d$, is associated with $\mathbf{B}_h$. It is also not difficult to show that each diagonal block matrix $\mathbf{P}_h$ is given by

$$\mathbf{P}_h = \mathbf{P}_{h,h-1}\mathbf{P}_{h-1,h-2}...\mathbf{P}_{2,1}\mathbf{P}_{1,d}\mathbf{P}_{d,d-1}...\mathbf{P}_{h+1,h}, \tag{23}$$

where $\mathbf{P}_{i,j}$ are those shown in (25) of [3].

Under $\mathbf{R}$, $\mathbf{B}_h$ is an acyclic group that can be treated as a Markov chain by itself with a transition probability matrix $\mathbf{P}_h$. Its limiting probability distribution can then be found by using (18)–(21). Actually the limiting probability distributions for these subgroups are not independent. When one is found the others can be computed readily. For instance, let

$$\hat{\mathbf{p}}_{(h)} = \{p_{(h)1}, p_{(h)2},..., p_{(h)N_h}\}^T, \quad h=1,2,...,d \tag{24}$$

denote the limiting probability distribution for the hth subgroup under the action of $\mathbf{R}$. Then one can show that

$$\hat{\mathbf{p}}_{(h+1)} = \mathbf{P}_{h+1,h}\,\hat{\mathbf{p}}_{(h)}, \quad h=1,2,...,d-1, \qquad \hat{\mathbf{p}}_{(1)} = \mathbf{P}_{1,d}\,\hat{\mathbf{p}}_{(d)}, \tag{25}$$

where $\mathbf{P}_{i,j}$ are the block matrices appearing in (25) of [3]. For the overall discussion of the periodic group as a whole it is convenient to introduce a set of d limiting probability vectors as follows:

$$\mathbf{p}_{(h)} = \{0,0,...,0,p_{(h)1},p_{(h)2},...,p_{(h)N_h},0,0,...,0\}^T \tag{26}$$

[4] In this section $\mathbf{P}$ refers to the transition probability matrix of a persistent group not that of the whole system. Similarly, N will stand for the number of cells in the group.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where $h=1,2,\ldots,d$. The hth vector in the set has nonzero components only at the positions for cells belonging to the hth subgroup.

Let us now describe the propagation of the probability distribution among the cells of this group during the evolution of the system. Let the system start from cells all in one subgroup, say $\mathbf{B}_h$. Then as the system evolves, it will occupy cells in $\mathbf{B}_{h+1}$, $\mathbf{B}_{h+2},\ldots,\mathbf{B}_d$, $\mathbf{B}_1,\ldots$successively, going through all the subgroups in a cyclic fashion. At each stage of the evolution the probability distribution among the cells of the subgroup that the system occupies can be computed by using the transition probability matrix. Eventually, the probability distribution approaches a limit in the following sense:

$$\lim_{\substack{n\to\infty \\ n=g \bmod d}} \zeta(n)=\mathbf{p}_{((g+h) \bmod d)} \tag{27}$$

Next consider a system that starts with a probability distribution covering cells that belong to different subgroups. Let $\zeta(0)$ be the initial probability vector. From $\zeta(0)$ we can compute the initial probability for each subgroup $\mathbf{B}_h$ and denote it by $\eta_{(h)}(0)$.

$$\eta_{(h)}(0)=\sum_{i\in\mathbf{B}_h}\zeta_i(0), \quad h=1,2,\ldots,d. \tag{28}$$

Let us now introduce a rectangular matrix $\mathbf{P}^{(\Omega)}$, a subgroup probability vector $\eta(0)$, and a permutation matrix $\mathbf{E}$ as follows:

$$\mathbf{P}^{(\Omega)}=(\mathbf{p}_{(1)},\mathbf{p}_{(2)},\ldots,\mathbf{p}_{(d)}), \quad \text{of order } N\times d,$$

$$\eta(0)=\{\eta_{(1)}(0),\eta_{(2)}(0),\ldots,\eta_{(d)}(0)\}^T, \quad \text{of order } d\times 1,$$

$$\mathbf{E}=\begin{bmatrix} 0 & 0 & . & 0 & 1 \\ 1 & 0 & . & 0 & 0 \\ 0 & 1 & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & 1 & 0 \end{bmatrix} \quad \text{of order } d\times d. \tag{29}$$

The limiting probability vector $\zeta(n)$ can then be written as

$$\lim_{\substack{n\to\infty \\ n=g \bmod d}} \zeta(n)=\mathbf{P}^{(\Omega)}\mathbf{E}^g\eta(0) \tag{30}$$

7 Determination of the Transient Groups

We now turn to transient cells. In this section we describe a procedure of dividing them into transient groups $\mathbf{B}_{G_p+1}$, $\mathbf{B}_{G_p+2},\ldots,\mathbf{B}_{G_p+G_t}$ such that the system can go from one group to a group with an equal or lower subscript designation but not to one with a higher designation. Let the total number of transient cells be N_t. First we construct a cell space $\mathbf{S}_1$ with $Y=1,2,\ldots,N_t$ as the cell designation by deleting from the original cell space $\mathbf{S}$ all the persistent cells. The relation between the Y cells and the retained Z cells is to be given by an array $S_1(Y)$.

$$S_1(Y)=Z. \tag{31}$$

Next we examine the image cells of each Y cell and delete those that are persistent cells. In this manner a new array $C_1(Y,i)$ is created where i ranges from 0 to $I_1(Y)$ with $I_1(Y)$ denoting the number of images remaining after deletion. By this construction we obtain a new cell space $\mathbf{S}_1$ of Y cells and the associated arrays $C_1(Y,i)$ and $I_1(Y)$. Thereafter we take the following steps:

(*i*) By (17) of [3] it is obvious that if a cell Y has its $I_1(Y)=0$, then it is a transient group of a single cell. Therefore, we sweep through the $\mathbf{S}_1$ cell space to single out all transient groups of this kind. These transient cells will each be assigned with group numbers G_p+1, $G_p+2,\ldots,G_p+h$, respectively. For example, if $Y=5$ is the second transient group discovered in this manner, then we have $G(S_1(5))=G_p+2$. We shall also assign the member number $N(g)$ for each group. In the present case $N(g)=1$ for $g=G_p+1$, $G_p+2,\ldots$, G_p+h. These groups will be called *discovered transient groups*.

(*ii*) Next, we construct from the old $\mathbf{S}_1$ cell space a new $\mathbf{S}_1$ cell space by deleting all the cells of the discovered transient groups. We also construct a new identification array $S_1(Y)=Z$, a new image cell array by also deleting all the image cells that are cells of the discovered transient groups, and a new image number array $I_1(Y)$.

(*iii*) On this new cell space we repeat Step (*i*).

(*iv*) By repeating (*i*), (*ii*), and (*iii*), sooner or later we arrive at an updated cell space such that $I_1(Y)$ for every Y in that space is nonzero. On that space we use the algorithm described in Section 4 in order to find the next transient group, the cells of which communicate among themselves. When such a transient group is discovered, all cells in the group are assigned the next group number and this group becomes a discovered transient group. The member number $N(g)$ for that group will also be assigned.

(*v*) We repeat the process of creating a new cell space by deleting the cells of this newly discovered transient group. We also delete these cells from the image cell list as before to create a new array $C_1(Y,i)$ and a new array $I_1(Y)$. After that we repeat the cycle with step (*i*).

In this manner we can determine all the transient groups and the normal form of $\mathbf{P}$ in the form of (17) of [3] is now completely in place.

8 Absorption Probability and Expected Absorption Time

Once we have determined the transient groups, we are in a position to compute the absorption probability of a transient cell j into a persistent cell i and the expected absorption time of a transient cell into persistent groups. Here our method is based on Theorems 14 and 15 of [3]. Needed is

$$\mathbf{N}=(\mathbf{I}-\mathbf{Q})^{-1} \tag{32}$$

$\mathbf{Q}$ is, of course, of a huge order. To find the inverse of such a matrix is a very difficult task. However, for our problem $\mathbf{Q}$ is usually a very sparse matrix and so will be $(\mathbf{I}-\mathbf{Q})$. Finding the inverse $\mathbf{N}$ can be facilitated by using sparse matrix techniques. Moreover, we note that when $\mathbf{P}$ is put in its normal form (17) of [3], $\mathbf{Q}$ is in the upper triangular block matrix form. This implies that $(\mathbf{I}-\mathbf{Q})$ and hence $\mathbf{N}$ will have the same form.

$$\mathbf{N}=\begin{bmatrix} \mathbf{N}_{1,1} & \mathbf{N}_{1,2} & . & \mathbf{N}_{1,m} \\ & \mathbf{N}_{2,2} & . & \mathbf{N}_{2,m} \\ & & . & . \\ & & & \mathbf{N}_{m,m} \end{bmatrix} \tag{33}$$

Moreover, it is easy to show that

$$\mathbf{N}_{i,i}=(\mathbf{I}-\mathbf{Q}_{k+i})^{-1}, \quad i=1,2,\ldots,m$$

$$\mathbf{N}_{i,j}=\sum_{h=i}^{j-1}\mathbf{N}_{i,h}\,\mathbf{T}_{k+h,k+j}\,\mathbf{N}_{j,j}, \qquad \begin{matrix} i=1,2,\ldots,m-1, \\ j=i+1,i+2,\ldots,m, \end{matrix} \tag{34}$$

where $\mathbf{Q}_{k+h}$ and $\mathbf{T}_{i,j}$ are block matrices shown in (17) of [3]. Also we have $k=G_p$ and $m=G_t$. With regard to notation, we shall use matrix $\mathbf{N}_{i,j}$ to denote the block matrix at the (i,j)th position in $\mathbf{N}$. The scalar N_{ij} is used to denote the (i,j)th element of $\mathbf{N}$. The block matrices may be computed in a sequential order $\mathbf{N}_{1,1}$, $\mathbf{N}_{2,2},\ldots$, $\mathbf{N}_{m,m}$, $\mathbf{N}_{1,2}$, $\mathbf{N}_{2,3}$, ..., $\mathbf{N}_{m-1,m}$, $\mathbf{N}_{1,3}$, ..., $\mathbf{N}_{1,m}$.

Once $\mathbf{N}$ is determined, we can compute the expected ab-

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

sorption time of a transient cell j into the persistent groups

$$\nu_j = \sum_{i=1}^{N_t} N_{ij} \tag{35}$$

where N_t is the number of the transient cells. Having $\mathbf{N}$, one can also then compute

$$[\alpha_{ij}] = \mathbf{T}\,\mathbf{N} \tag{36}$$

to find the absorption probability of a transient cell j into a persistent cell i. $[\alpha_{ij}]$ is the most important set of global results. For each j a nonzero α_{ij} means that the persistent group to which the cell i belongs is a domicile of j. Thus, by examining all α_{ij} for a given j we can locate all the domiciles of j. Moreover, the group absorption probability of every transient cell j into one of its domiciles can then be determined by (5).

9 Variations of the Basic Scheme and Possible Improvements

In Sections 3–8 we have presented the basic algorithm for analyzing Markov chains. It is offered in such a manner as to present the ideas in the simplest way without being encumbered by the question of efficiency. In this section we discuss various steps one can take to improve either the storage requirements or the computing time.

Sparse Matrix Techniques. Whenever the inversion of a large matrix or the product of two large matrices is required, it is advantageous to employ sparse matrix techniques, Duff [7].

Iteration Procedure. To compute the limiting probability of an acyclic group or a subgroup of a periodic persistent group, instead of solving (21), it is sometimes more efficient to use the iteration scheme

$$\mathbf{p} = \lim_{n\to\infty} \zeta(n), \quad \zeta(n+1) = \mathbf{P}\zeta(n), \tag{37}$$

starting with an *arbitrary* vector $\zeta(0)$. The convergence of $\zeta(n)$ to $\mathbf{p}$ as $n\to\infty$ is assured.

The Smallest Subgroup. With regard to finding $\hat{\mathbf{p}}_{(h)}$, $h=1,2,\ldots,d$, of the subgroups of a periodic group, because of (25) we should first pick a subgroup that has the least members and compute its limiting probability. The limiting probability distributions of the other subgroups are then easily generated through (25).

Transient Cells as One Group. Of all the parts of the algorithm the most time-consuming part is the evaluation of the matrix $\mathbf{N} = (\mathbf{I}-\mathbf{Q})^{-1}$. In Sections 7 and 8 one is assumed to find the transient groups first and then to find $\mathbf{N}$ by (33) and (34). In computing $\mathbf{N}$ one will invariably use sparse matrix techniques. The one devised by Duff [7] has a preliminary part of preordering the matrix which is, in essence, equivalent to finding the transient groups. When a code of that kind is used for inverting $(\mathbf{I}-\mathbf{Q})$, there is no need to locate the transient groups separately.

Not Storing the Complete N. The usual sparse matrix method of inverting $(\mathbf{I}-\mathbf{Q})$ is to solve for $\mathbf{x}^{(i)}$ from

$$(\mathbf{I}-\mathbf{Q})\mathbf{x}^{(i)} = \mathbf{b}^{(i)}. \tag{38}$$

By taking

$$\mathbf{b}^{(i)} = \{0,0,\ldots,0,{}_{i\text{th}}1,0,\ldots,0\}^T, \quad i=1,2,\ldots,N_t, \tag{39}$$

and solving for $\mathbf{x}^{(i)}$, one obtains

$$\mathbf{N} = (\mathbf{I}-\mathbf{Q})^{-1} = \{\mathbf{x}^{(1)},\mathbf{x}^{(2)},\ldots,\mathbf{x}^{(N_t)}\}. \tag{40}$$

By (36) we have

$$[\alpha_{ij}] = \mathbf{T}\,\mathbf{N} = \{\mathbf{T}\mathbf{x}^{(1)},\mathbf{T}\mathbf{x}^{(2)},\ldots,\mathbf{T}\mathbf{x}^{(N_t)}\}. \tag{41}$$

This leads to a simplification in computing. We can first compute $\mathbf{x}^{(1)}$ by (38). By Theorem 14 of [3] the sum of the elements of $\mathbf{x}^{(1)}$ gives us the expected absorption time of this first transient cell to be absorbed into the persistent groups. Next we compute $\mathbf{T}\mathbf{x}^{(1)}$ which is the first column $[\alpha_{i1}]$ of $[\alpha_{ij}]$. This column gives us the probabilities of the transient cell No. 1 being absorbed into various persistent cells[5]. Once α_{i1} values have been found, there is no further need to store $\mathbf{x}^{(1)}$. We can then go to (38) and solve for $\mathbf{x}^{(2)}$. This vector and $\mathbf{T}\mathbf{x}^{(2)}$ then provide us with the expected absorption time and the absorption probabilities of the second transient cell into the persistent cells. The process is then repeated for all the transient cells. The preceding discussion shows that it is unnecessary to store the entire matrix $\mathbf{N}$. This is a very important observation, because while $(\mathbf{I}-\mathbf{Q})$ is in general very sparse, its inverse may not be so. Therefore, a complete storage of $\mathbf{N}$ is undesirable.

A Partition of Q. Sometimes it is advantageous to partition $\mathbf{Q}$ in such a way so that the matrix $\mathbf{P}$ of (37) of [3] takes the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_p & \mathbf{T} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_p & \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_t \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_t \end{bmatrix}. \tag{42}$$

Cells associated with the columns occupied by $\mathbf{T}_1$ will be called Type 1 transient cells. Each of them will only have persistent cells as its images and, moreover, all its image cells are in *one* persistent group only. Cells associated with the columns occupied by $\mathbf{T}_2$ will be called Type 2 transient cells. Again, each of them will have only persistent cells as its images but the image cells occupy more than one persistent group. Cells associated with $\mathbf{Q}_t$ will be called Type 3 transient cells; they will have transient cells as well as persistent cells as their images. With this partition of $\mathbf{P}$ one can show that

$$\mathbf{N} = (\mathbf{I}-\mathbf{Q})^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & -\mathbf{Q}_1 \\ \mathbf{0} & \mathbf{I} & -\mathbf{Q}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{I}-\mathbf{Q}_t \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{Q}_1\hat{\mathbf{N}} \\ \mathbf{0} & \mathbf{I} & \mathbf{Q}_2\hat{\mathbf{N}} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{N}} \end{bmatrix} \tag{43}$$

$$\mathbf{T}\,\mathbf{N} = (\mathbf{T}_1,\mathbf{T}_2,\hat{\mathbf{T}}\hat{\mathbf{N}}) \tag{44}$$

where

$$\hat{\mathbf{T}} = \mathbf{T}_1\mathbf{Q}_1 + \mathbf{T}_2\mathbf{Q}_2 + \mathbf{T}_t$$
$$\hat{\mathbf{N}} = (\mathbf{I}-\mathbf{Q}_t)^{-1} \tag{45}$$

When there are a very large number of Type 1 and Type 2 transient cells, using (43) and (44) to compute $\mathbf{N}$ and $\mathbf{TN}$ leads to better efficiency.

Merging of Cells. Another improvement of efficiency can be done when one is interested in the determination of absorption probabilities of transient cells into persistent groups, but not interested in the expected absorption times or the eventual probability distributions among the persistent cells. In that case we can merge cells. First, the set of all cells in a persistent group can be taken to be one amalgamated absorbing cell. Thus, k persistent groups are replaced by k absorbing cells of various size, and the transition probability

[5] A simple check for the accuracy of inverting is provided here by the requirement $\sum_i \alpha_{ij} = 1$.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

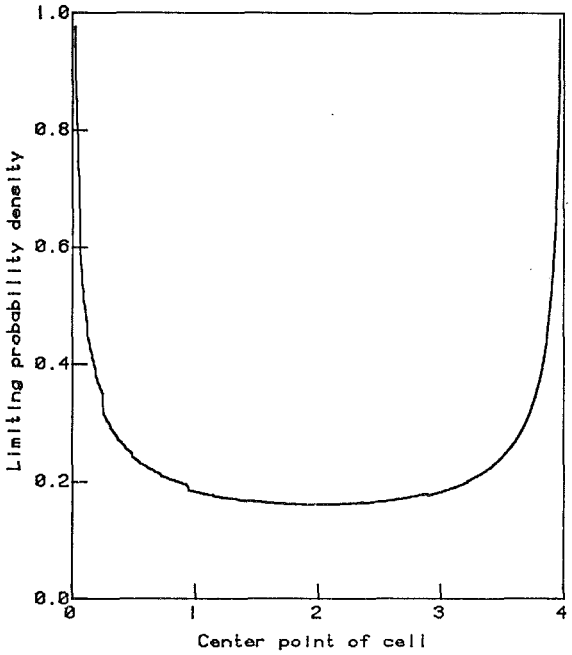


Fig. 1 The limiting probability density of the long-term chaotic motion for the population dynamics model (46) with $s=4$ and using 4000 regular cells

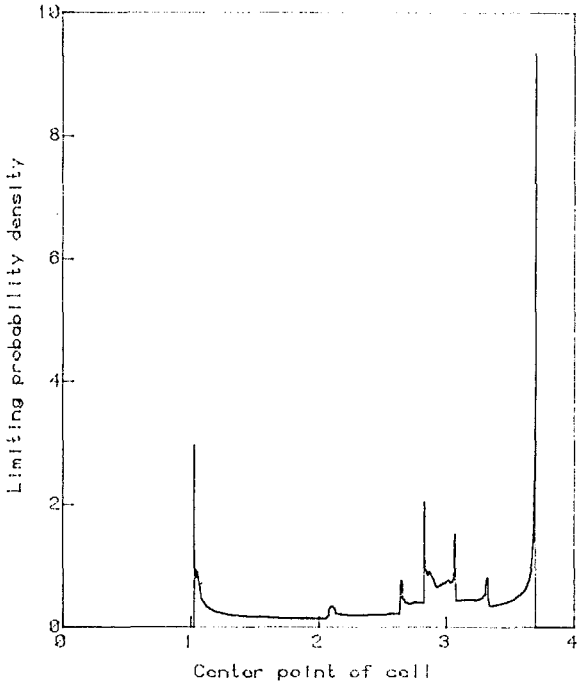


Fig. 2 The limiting probability density of the long-term chaotic motion for the population dynamics model (46) with $s=3.7$ and using 4000 regular cells

matrix **P** is to be adjusted accordingly. Secondly, if a transient cell is known to be eventually mapped 100 percent into a persistent group, then that cell can be amalgamated into that group and the total number of cells to be handled is reduced by one. This second idea of merging a transient cell to a persistent group has been tried out as a basis of devising an iterative process for analyzing generalized cell mappings and found promising. However, the approach is not yet fully explored and any detailed discussion is premature at this time.

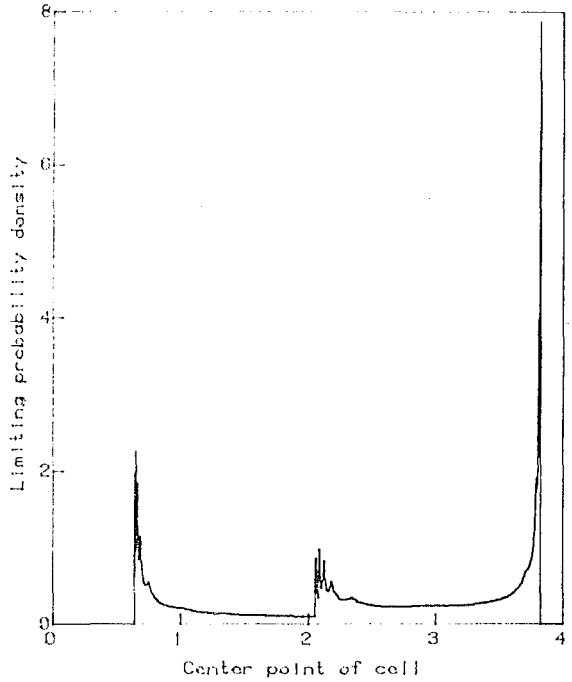


Fig. 3 The limiting probability density of the long-term chaotic motion for the population dynamics model (46) with $s=3.825$ and using 4000 regular cells

10 Examples of Application

In this section we demonstrate the viability of the method by applying it to a one and two-dimensional point mapping system. The one-dimensional system to be examined is the same one of population dynamics model studied in [3],

$$x(n+1)=s\,x(n)[1-x(n)/4] \tag{46}$$

except that in [3] only a very small number of cells were used to introduce the basic idea of generalized cell mappings. Consider $s=4$. For this case the point mapping motion is chaotic, wandering over the range $0 \leqq x \leqq 4$, if the motion starts inside this range. Analytically, it has been established that the long-term occupancy rate of this chaotic motion is governed by, [8],

$$\frac{1}{\pi\sqrt{x(4-x)}} \tag{47}$$

When the generalized cell mapping method is applied to this problem and 4000 regular cells are used to cover the range 0–4 with a sink cell covering the outside, the result is that all the 4000 regular cells belong to a huge acyclic persistent group. One can then compute the limiting probability distribution of this group. As shown in Fig. 1 it duplicates excellently the analytical result of (47)[6].

Next, we consider the cases where $s=3.7$ and 3.825, respectively. The long-term motion of the point mapping are again chaotic but covering only a portion of the range 0–4. Using the generalized cell mapping method and using regular cells to cover 0–4, we find again that in each case the long-term chaotic motion is replaced by a huge acyclic group with all other regular cells being transient. For these values of s there does not seem to exist any analytical result predicting the long-term occupancy rate of the chaotic motion of the point mapping. With the generalized cell mapping one can readily compute the limiting probability of the acyclic groups. They are shown in Figs. 2 and 3. In each case 4000 regular cells are

[6] Plotted in Figs. 1-3 is the limiting probability density which is the limiting probability of each cell divided by the cell size.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

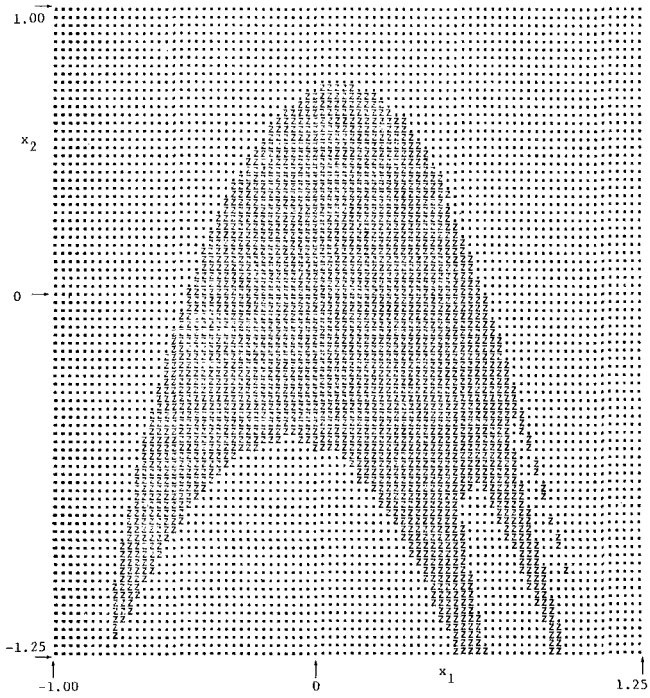


Fig. 4 For system (48). Cells denoted by z are absorbed 100 percent to the acyclic persistent group B_2 at the origin. Phase plane covered: $-1.00 \leqq x_1 < 1.25$, $-1.25 \leqq x_2 < 1.00$. $h_1 = h_2 = 0.028125$. Total number of regular cells = 6400.

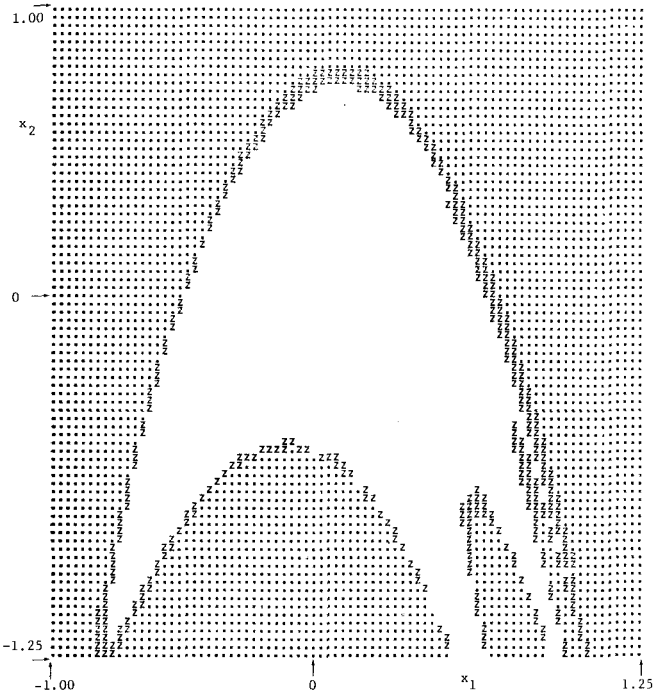


Fig. 5 For system (48). Cells denoted by z are absorbed to the acyclic persistent group B_2 at the origin with a probability in the range 95 percent $\leqq A_1(Z,2) < 100$ percent. Phase plane covered: $-1.00 \leqq x_1 < 1.25$, $-1.25 \leqq x_2 < 1.00$. $h_1 = h_2 = 0.028125$. Total number of regular cells = 6400.

used to cover the range 0–4. When these results are compared with the results obtained by using 1000 or 2000 cells, one finds that the general shape of the curve in each case is about the same but spikes are less pronounced when more cells are used. One can also carry out the point mapping directly, say a million mapping steps to establish the chaotic motion, and

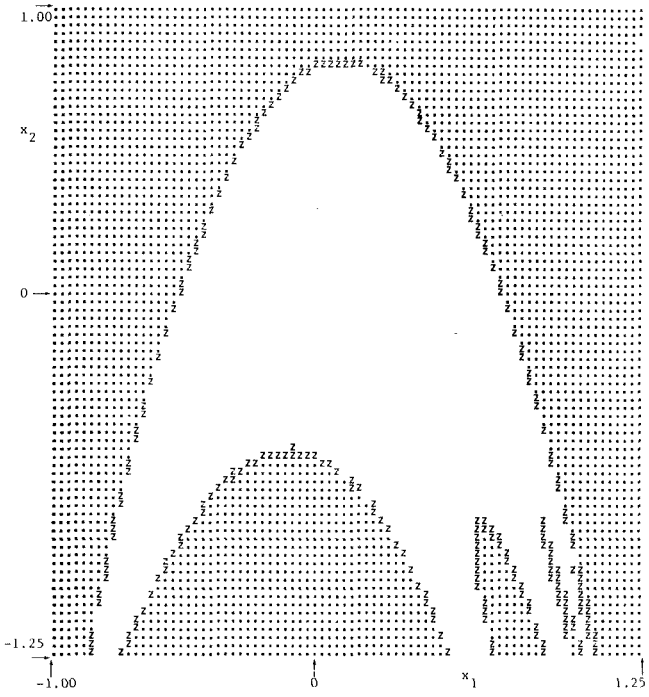


Fig. 6 For system (48). Cells denoted by z are absorbed to the acyclic persistent group B_2 at the origin with a probability in the range 50 percent $\leqq A_1(Z,2) < 95$ percent. Phase plane covered: $-1.00 \leqq x_1 < 1.25$, $-1.25 \leqq x_2 < 1.00$. $h_1 = h_2 = 0.028125$. Total number of regular cells = 6400.

then map several million more times to compute the occupancy rate along the x-axis by the chaotic motion. The limiting probability distributions computed by the cell mapping agree very well with the direct count results except that the direct count results have more pronounced spikes.

When the foregoing results are examined, one cannot help but appreciate that when there is a chaotic motion, the use of the limiting probability distribution of a persistent group is perhaps a more sensible way to describe the motion than trying to follow the chaotic motion of the point mapping endlessly.

The second problem used to test the method is the two-dimensional map:

$$x_1(n+1) = 0.9\, x_2(n) + 1.81\, x_1^2(n)$$
$$x_2(n+1) = -0.9\, x_1(n). \tag{48}$$

This system has an asymptotically stable spiral point at the origin and a saddle point at $(1, -0.9)$. The problem is picked because we know rather completely its nonlinear global behavior. The system was also treated in [2] by using the simple cell-mapping method. The domain of attraction for the stable spiral point at the origin is given by Fig. 2 of [2]. Concerning that figure, one could however raise a question that a cell near the boundary will be assigned inside or outside the domain depending on whether the center point of the cell is inside or out. Thus, the crudeness of the simple cell mapping probably does not allow us to interpret the behavior of the boundary cells in a definitive way. Also, near the saddle point there is a sequence of "unstable" finger-shaped regions, with the fingers becoming narrower and narrower as they approach the separatrix. Again, with the crude simple cell mapping one cannot hope to find these fingers unless one uses very small cells.

By using the generalized cell mapping one has a more natural way of describing the properties of the cells on the boundary of a domain of attraction or in a region where two domains of attraction intertwine. That is to use the various

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

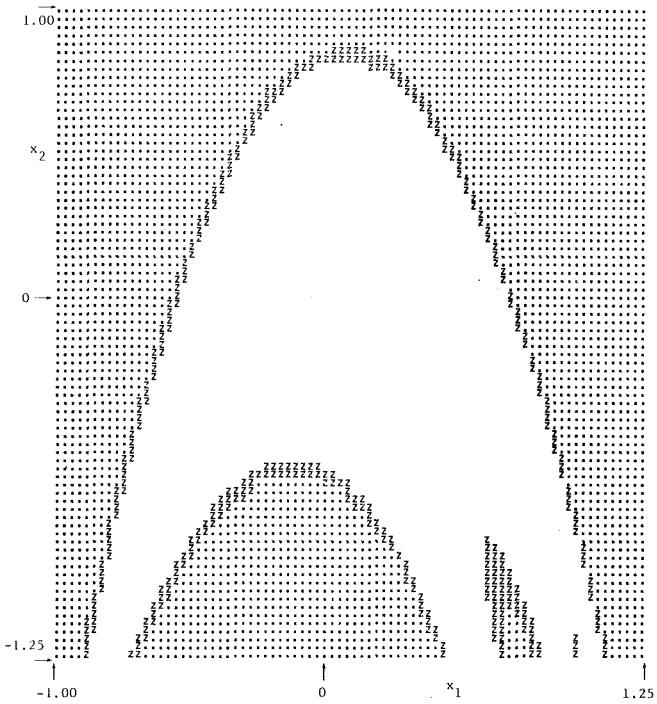


Fig. 7 For system (48). Cells denoted by z are absorbed to the acyclic persistent group B_2 at the origin with a probability in the range $0 < A_1(Z,2) < 50$ percent. Phase plane covered: $-1.00 \leqq x_1 < 1.25$, $-1.25 \leqq x_2 < 1.00$. $h_1 = h_2 = 0.028125$. Total number of regular cells = 6400.

absorption probabilities of a cell into its domiciles. Thus, for a cell near the boundary we can find what fraction of the trajectories from this cell will go to one persistent group and the remaining will go to another persistent group. In a similar way, for a cell traversed in a very complicated way by two domains of attraction of the point mapping, we can find the two domiciles of this cell and determine the respective absorption probability of this cell to each. Again this probabilistic description seems to be a much more sensible way to describe the situation rather than trying to chase the intertwining pattern of the two domains of attraction finer and finer.

Sampling Method of Generating the Cell Mappings. To create a generalized cell mapping for a given two-dimensional point mapping, one question arises immediately, which is how to compute **P**. There are different ways to achieve this, but the one that is simple and robust is the "*Sampling Method.*" To find the image cells of a cell Z and the associated probabilities, one simply takes M number of points, say uniformly distributed, inside the cell Z and computes the point mapping images of these M points. If M_1 image points lie in cell $Z^{(1)}$, M_2 points lie in cell $Z^{(2)}$,..., and M_m points in cell $Z^{(m)}$, then we assign

$$I(Z) = m, \quad C(Z,i) = Z^{(i)}, \quad P(Z,i) = M_i/M, \quad i = 1,2,\ldots,I(Z) \tag{49}$$

This is an extremely simple and effective method, usable for any nonlinearity and can be extended to point mappings of any dimension.

Using this sampling method and using 100 sampling points in each cell, a generalized cell mapping is created for (48) and the Markov chain so generated is analyzed by the present algorithm. One finds that there is an acyclic persistent group $\mathbf{B}_2$ situated at the origin. The global behavior of other cells is shown in Figs. 4–7. Figure 4 shows by the symbol z all the cells that have only one domicile which is $\mathbf{B}_2$. Other cells in Fig. 4 will have two domiciles or one domicile which is not $\mathbf{B}_2$. Figure 5 shows a thin strip of cells that are absorbed into $\mathbf{B}_2$ with a probability in the range 95–100 percent. Figure 6 shows cells that are absorbed into $\mathbf{B}_2$ with a probability in the range 50–95 percent. Similarly, Fig. 7 shows cells in the absorption probability range 0–50 percent. In other words, cells represented by symbol • in Fig. 7 are those that have only the sink cell as its only domicile. Here the sink cell covers the phase plane outside the rectangle shown.

11 Concluding Remarks

In this paper we have presented an algorithm for analyzing generalized cell mappings. We believe this approach represents a new departure for global analysis of nonlinear systems. Using a probabilistic description for the state, many features of strongly nonlinear systems involving chaotic behavior and intertwining domains of attraction can be dealt with in a more satisfactory manner. Also, the feasibility of using the method is not made any less by the presence of strong nonlinearity. Moreover, using a cell-structured state space, instead of a continuum state space, the theory lends itself most naturally to computer implementation. One could expect the power and potential of this approach to go hand in hand with the development of computer capability. Of course, we should also mention here that this is only the first paper dealing with generalized cell mappings involving very large number of cells. We anticipate improvements and innovations on the methods of analyzing such cell mappings in the future.

Acknowledgment

This material is based on work supported by the National Science Foundation under Grant No. MEA-8019274.

References

1 Hsu, C. S., "A Theory of Cell-to-Cell Mapping Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 931–939.

2 Hsu, C. S., and Guttalu, R. S., "An Unravelling Algorithm for Global Analysis of Dynamical Systems: An Application of Cell-to-Cell Mappings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 940–948.

3 Hsu, C. S., "A Generalized Theory of Cell-to-Cell Mapping for Nonlinear Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 634–642.

4 Isaacson, D. L., and Madsen, R. W., *Markov Chains: Theory and Applications*, Wiley, New York, 1976.

5 Chung, K. L., *Markov Chains With Stationary Transition Probabilities*, 2nd Ed., Springer-Verlag, New York, 1967.

6 Romanovsky, V. L., *Discrete Markov Chains*, Wolers-Noordhoff Publishing, Groningen, The Netherlands, 1970.

7 Duff, I. S., "MA28—A Set of FORTRAN Subroutines for Sparse Unsymmetric Linear Equations," AERE-R.8730 Report, Harwell, England.

8 Lieberman, M. A., "Regular and Chaotic Motion in Classical Dynamics," Private lecture notes, kindly provided to the first author of this paper by Professor Lieberman.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Probabilistic Theory of Nonlinear Dynamical Systems Based on the Cell State Space Concept

C. S. Hsu
Professor of Applied Mechanics,
Department of Mechanical Engineering,
University of California,
Berkeley, Calif. 94720
Fellow ASME

Developed in the paper is a probabilistic theory for nonlinear dynamical systems. The theory is based on discretizing the state space into a cell structure and using the cell probability functions to describe the state of a system. Although the dynamical system may be highly nonlinear the probabilistic formulation always leads to a set of linear ordinary differential equations. The evolution of the probability distribution among the cells can then be studied by applying the theory of Markov processes to this set of equations. It is believed that this development possibly offers a new approach to the global analysis of nonlinear systems.

1 Introduction

If one is interested in the long-term behavior of a highly nonlinear system governed by a set of ordinary differential equations and if one is given a specific set of initial conditions, then the most commonly used method is perhaps still the straightforward numerical integration. However, if one is interested in the quantitative determination of the general global behavior of such a system, this straightforward approach would require extensive numerical integration over a very large number of sets of initial conditions and the method is usually no longer viable. For such a task it is desirable to search for other more effective approaches. In this paper we present a new approach of this kind.

The basic idea behind this new approach is not to consider various motions individually but rather to consider an ensemble of a very large number of motions or trajectories that start with different sets of initial conditions. One examines the global behavior of the system by determining how, as time increases, the trajectories are distributed probabilistically in the state space if the probability distribution of the initial conditions is known. The theory to be presented uses two key ingredients. One is to use the cell-structured state space instead of the usual state space of a continuum. The philosophical and physical bases for introducing the cell state space and the usefulness of the concept has been discussed in [1-3]. Using that concept, the N-dimensional state space is regarded as a collection of N-dimensional cells, with each cell considered as an entity. Without any loss of generality each cell can be identified with an N-tuple of integers $\mathbf{Z}(Z_1, Z_2, \ldots, Z_N)$. $\mathbf{Z}$ will be called a *cell vector*, which is always integer valued.

The second ingredient is to use a probabilistic description of the state of the system. Consider a very large ensemble of motions and let $\zeta(\mathbf{Z},t)$ be the probability of finding the system at cell $\mathbf{Z}$ at time t. Evidently, we have

$$0 \leqq \zeta(\mathbf{Z},t) \leqq 1, \tag{1}$$

$$\sum_{\mathbf{Z}} \zeta(\mathbf{Z},t) = 1, \tag{2}$$

where the sum ranges over all the cells. For convenience $\zeta(\mathbf{Z},t)$ will be called the *cell probability function*.

Having introduced the cell state space and the cell probability function, we are now in a position to discuss the new theory. This will be done in Sections 2 and 3 where the theory is developed for one and two-dimensional cases. Extension to higher dimensional systems is obvious. In Sections 4-6 we give some general discussions on the nature of the differential equations derived in Sections 2 and 3 and on certain properties of the solutions of these equations, especially their bearing on the global behavior of the systems.

2 One-Dimensional Systems

In order to discuss the basic idea underpinning the present theory we shall first treat the case of one-dimensional systems. Consider a system governed by

$$\frac{dx}{dt} = F(x,t). \tag{3}$$

Let the axis of the state variable x be divided into a cell structure. A regular cell Z with a range $Z' \leqq Z \leqq Z''$ is defined to be one that contains all x satisfying

$$(Z - \frac{1}{2})h \leqq x < (Z + \frac{1}{2})h \tag{4}$$

where h is the cell size.

Sink Cells. Following the idea discussed in [2, 3] we introduce two sink cells. We assume that our interest in the behavior of the system is confined to the range from (Z'

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14-19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January, 1982. Paper No. 82-WA/APM-22.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

evolution of the cell probability functions of all the cells. These equations can be put into the following form:

$$\frac{d}{dt}\begin{bmatrix}\zeta(Z_{LS})\\ \zeta(Z')\\ \zeta(Z'+1)\\ .\\ \zeta(Z-1)\\ \zeta(Z)\\ \zeta(Z+1)\\ .\\ \zeta(Z''-1)\\ \zeta(Z'')\\ \zeta(Z_{US})\end{bmatrix} = \begin{bmatrix}0 & x & 0 & . & 0 & 0 & 0 & . & 0 & 0 & 0\\ 0 & x & x & . & 0 & 0 & 0 & . & 0 & 0 & 0\\ 0 & x & x & . & 0 & 0 & 0 & . & 0 & 0 & 0\\ . & . & . & . & . & . & . & . & . & . & .\\ 0 & 0 & 0 & . & x & x & 0 & . & 0 & 0 & 0\\ 0 & 0 & 0 & . & x & x & x & . & 0 & 0 & 0\\ 0 & 0 & 0 & . & 0 & x & x & . & 0 & 0 & 0\\ . & . & . & . & . & . & . & . & . & . & .\\ 0 & 0 & 0 & . & 0 & 0 & 0 & . & x & x & 0\\ 0 & 0 & 0 & . & 0 & 0 & 0 & . & x & x & 0\\ 0 & 0 & 0 & . & 0 & 0 & 0 & . & 0 & x & 0\end{bmatrix}\begin{bmatrix}\zeta(Z_{LS})\\ \zeta(Z')\\ \zeta(Z'+1)\\ .\\ \zeta(Z-1)\\ \zeta(Z)\\ \zeta(Z+1)\\ .\\ \zeta(Z''-1)\\ \zeta(Z'')\\ \zeta(Z_{US})\end{bmatrix} \tag{20}$$

where most of the elements of the square matrix are zero and the only possibly nonzero elements are indicated by the symbol x. For convenience the argument t in $\zeta(Z,t)$ has been deleted in this equation. This will also be done in the remainder of this paper unless a specific identification with t is necessary or desirable. Let ζ be a *cell probability vector* whose components are $\zeta(Z,t)$, $Z=Z_{LS}, Z', \ldots, Z'', Z_{US}$. Then we may write (20) as

$$\frac{d\zeta}{dt}=\mathbf{Q}(t)\zeta \tag{21}$$

where $\mathbf{Q}(t)$ will be called the *cell probability flow-rate matrix*, or *rate matrix* in abbreviated form. In this manner the dynamic behavior of the system is expressed in the form of an evolution equation of the cell probability vector.

Properties of the Rate Matrix Q(*t*). It is readily seen that the rate matrix has the following properties:

1. The first and the last columns associated with the sink cells are zero columns.
2. Each column has at most three nonzero elements.
3. Each row has at most three nonzero elements.
4. For the jth column, q_{ij} $(i\neq j)$ represents the flow rate of probability from cell j to cell i, and q_{jj} represents the rate of loss of probability of cell j to the two neighboring cells. The elements q_{ij} $(i\neq j)$ are always non-negative and q_{jj} is always nonpositive. Moreover,

$$q_{jj}=-\sum_{\substack{i\\ i\neq j}} q_{ij} \quad \text{or} \quad \sum_i q_{ij}=0 \tag{22}$$

i.e., each column has a zero sum.

3 Two-Dimensional Systems

Next we consider a two-dimensional system governed by

$$\frac{d\mathbf{x}}{dt}=\mathbf{F}(\mathbf{x},t)=\mathbf{F}(x_1,x_2,t) \tag{23}$$

where $\mathbf{x}(x_1,x_2)$ is a two-dimensional state vector and $\mathbf{F}$ is a two-dimensional vector function. Let the state space again be given a cell structure. We assume that our interest in the system is confined to

$$(Z_i'-\frac{1}{2})h_i \leqq x_i < (Z_i''+\frac{1}{2})h_i, \quad i=1,2 \tag{24}$$

where h_1 and h_2 are the cell sizes in the x_1 and x_2-directions. A regular cell $\mathbf{Z}$ with components Z_i $(Z_i' \leqq Z_i \leqq Z_i'', i=1,2)$ is defined to be one that contains all $\mathbf{x}$ satisfying

$$(Z_i-\frac{1}{2})h_i \leqq x_i < (Z_i+\frac{1}{2})h_i, \quad i=1,2. \tag{25}$$

Sink Cell. All $\mathbf{x}$-values outside the region of interest (24) are taken to belong to the *sink cell* which will be denoted by $\mathbf{Z}_S$. The cell space so defined now consists of the sink cell $\mathbf{Z}_S$ and $(Z_2''-Z_2'+1)(Z_1''-Z_1'+1)$ regular cells.

Let $\zeta(\mathbf{Z})$ again be the cell probability function for cell $\mathbf{Z}$ at time t. Now we examine the evolution of $\zeta(\mathbf{Z})$. Three kinds of cells will be considered separately.

Interior Cells. Consider an interior cell $\mathbf{Z}$ which has four edges at $x_1=x_1^-(\mathbf{Z}), x_1=x_1^+(\mathbf{Z}), x_2=x_2^-(\mathbf{Z})$, and $x_2=x_2^+(\mathbf{Z})$, or in an abbreviated notation at x_1^-, x_1^+, x_2^-, and x_2^+. We adopt the following notation for the neighboring cells of $\mathbf{Z}$: $\mathbf{Z}+\mathbf{e}_1$ for the cell to the right, $\mathbf{Z}-\mathbf{e}_1$ the cell to the left, $\mathbf{Z}+\mathbf{e}_2$ the cell above, and $\mathbf{Z}-\mathbf{e}_2$ the cell below; See Fig. 1. Again, let M be the number of trajectories in the ensemble and let $M(\mathbf{Z})$ be the number of trajectories momentarily residing in cell $\mathbf{Z}$ at time t. Again the cell probability function $\zeta(\mathbf{Z}, t)$ which is equal to $M(\mathbf{Z})/M$ will be assumed uniformly distributed in cell $\mathbf{Z}$.

Consider now the flow of trajectories through the left edge at $x_1=x_1^-$. If $F_1(x_1^-, x_2, t)>0$, there are trajectories entering into cell $\mathbf{Z}$ from cell $\mathbf{Z}-\mathbf{e}_1$. If $F_1(x_1^-,x_2,t)<0$, there are trajectories leaving cell $\mathbf{Z}$ going into cell $\mathbf{Z}-\mathbf{e}_1$. To find the net crossing into cell $\mathbf{Z}$ through the left edge we need to integrate along the edge from x_2^- to x_2^+. For convenience we introduce a switching function $S(\sigma; a, b)$ such that

$$S(\sigma;a,b) = \begin{cases} a & \text{if} \quad \sigma<0\\ 0 & \text{if} \quad \sigma=0\\ b & \text{if} \quad \sigma>0\end{cases} \tag{26}$$

The number of trajectories crossing the left edge into cell $\mathbf{Z}$ during an infinitesimal time interval dt is given by

$$\frac{dt}{h_1h_2}\int_{x_2^-}^{x_2^+} F_1(x_1^-,x_2)S(F_1(x_1^-,x_2); M(\mathbf{Z}),M(\mathbf{Z}-\mathbf{e}_1))\,dx_2. \tag{27}$$

Here entering into cell $\mathbf{Z}$ is counted as positive and leaving negative. Also in (27) and hereafter we omit the writing of t in the expression for $\mathbf{F}$ and its components when it is not absolutely necessary.

Similarly the numbers of trajectories crossing the right edge, the bottom edge, and the top edge into cell $\mathbf{Z}$ during dt are, respectively:

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$-1/2)h$ to $(Z''+1/2)h$. A lower sink cell, designated as Z_{LS}, will cover all x values less than $(Z'-1/2)h$ and an upper sink cell, designated as Z_{US}, will cover all x values equal or greater than $(Z''+1/2)h$. This completes the process of defining the cell space which is seen to consist of two sink cells and $(Z''-Z'+1)$ regular cells. The total number of the cells in the cell space will be denoted by N_c which is equal to $(Z''-Z'+3)$ for the present case.

Let $\zeta(Z,t)$ be the cell probability function for cell Z at time t as defined in Section 1. It satisfies of course (1) and (2) where the summation over Z in (2) will cover the two sink cells and all the regular cells.

Consider now a typical interior cell Z. Let $x^-(Z)$ and $x^+(Z)$ be the lower and upper ends of cell Z

$$x^-(Z)=(Z-\frac{1}{2})h, \quad x^+(Z)=(Z+\frac{1}{2})h. \tag{5}$$

Consider a very large ensemble of M trajectories. Let M_z denote the number of trajectories residing in cell Z at time t. Let us now assume that the distribution of the residence of these trajectories within the cell Z is uniform. Let $F(x^+(Z),t)>0$. During an infinitesimal interval dt there will be

$$\frac{1}{h}(dx)\Big|_{x^+(Z)} M_Z \tag{6}$$

number of trajectories which move across $x^+(Z)$ and hence are lost to cell Z. In (6), dx evaluated at $x^+(Z)$ is equal to

$$(dx)\Big|_{x^+(Z)} = F(x^+(Z),t)\,dt \tag{7}$$

If $F(x^+(Z),t)<0$, then there will be

$$-\frac{1}{h}(dx)\Big|_{x^+(Z)} M_{Z+1} \tag{8}$$

number of trajectories gained by cell Z due to the trajectories coming into cell Z from cell $(Z+1)$. Similarly, if $F(x^-(Z),t)>0$, then during dt there will be a gain in trajectories for cell Z equal to

$$\frac{1}{h}(dx)\Big|_{x^-(Z)} M_{Z-1} \tag{9}$$

and if $F(x^-(Z),t)<0$ there will be a loss for cell Z equal to

$$\frac{1}{h}(dx)\Big|_{x^-(Z)} M_Z \tag{10}$$

where in (9) and (10)

$$(dx)\Big|_{x^-(Z)} = F(x^-(Z),t)\,dt. \tag{11}$$

Expressed in terms of cell probability function one obtains

$$\frac{d}{dt}\zeta(Z,t)=\frac{1}{h}\{F(x^-(Z),t)A-F(x^+(Z),t)B\}, \tag{12}$$

where

$$A=\begin{cases}\zeta(Z,t) & \text{if } F(x^-(Z),t)<0\\ 0 & \text{if } F(x^-(Z),t)=0\\ \zeta(Z-1,t) & \text{if } F(x^-(Z),t)>0\end{cases}$$
$$B=\begin{cases}\zeta(Z+1,t) & \text{if } F(x^+(Z),t)<0\\ 0 & \text{if } F(x^+(Z),t)=0\\ \zeta(Z,t) & \text{if } F(x^+(Z),t)>0.\end{cases} \tag{13}$$

Next we consider the sink cells and the boundary regular cells. Here we take the position that once a trajectory moves into the sink cells we have no further interest in the future evolution of that trajectory. Thus, the sink cells are absorbing cells. The rules of evolution for them are:

$$\frac{d}{dt}\zeta(Z_{LS},t)=0 \quad \text{if} \quad F(x^-(Z'),t)\geqq 0$$
$$=-\frac{1}{h}F(x^-(Z'),t)\,\zeta(Z',t) \quad \text{if} \quad F(x^-(z'),t)<0, \tag{14}$$

$$\frac{d}{dt}\zeta(Z_{US},t)=0 \quad \text{if} \quad F(x^+(Z''),t)\leqq 0$$
$$=\frac{1}{h}F(x^+(Z''),t)\,\zeta(Z'',t) \quad \text{if} \quad F(x^+(Z''),t)>0. \tag{15}$$

In the same spirit the rules of evolution for the boundary cells Z' and Z'' are:

$$\frac{d}{dt}\zeta(Z',t)=\frac{1}{h}\{F(x^-(Z'),t)C'-F(x^+(Z'),t)D'\} \tag{16}$$

where

$$C'=\begin{cases}\zeta(Z',t) & \text{if } F(x^-(Z'),t)<0\\ 0 & \text{if } F(x^-(Z'),t)\geqq 0\end{cases}$$
$$D'=\begin{cases}\zeta(Z'+1,t) & \text{if } F(x^+(Z'),t)<0\\ 0 & \text{if } F(x^+(Z'),t)=0\\ \zeta(Z',t) & \text{if } F(x^+(Z'),t)>0\end{cases} \tag{17}$$

and

$$\frac{d}{dt}\zeta(Z'',t)=\frac{1}{h}\{F(x^-(Z''),t)C''-F(x^+(Z''),t)D''\} \tag{18}$$

where

$$C''=\begin{cases}\zeta(Z'',t) & \text{if } F(x^-(Z''),t)<0\\ 0 & \text{if } F(x^-(Z''),t)=0\\ \zeta(Z''-1,t) & \text{if } F(x^-(Z''),t)>0\end{cases}$$
$$D''=\begin{cases}0 & \text{if } F(x^+(Z''),t)\leqq 0\\ \zeta(Z'',t) & \text{if } F(x^+(Z''),t)>0.\end{cases} \tag{19}$$

Equations (12), (14)-(16), and (18) now govern the

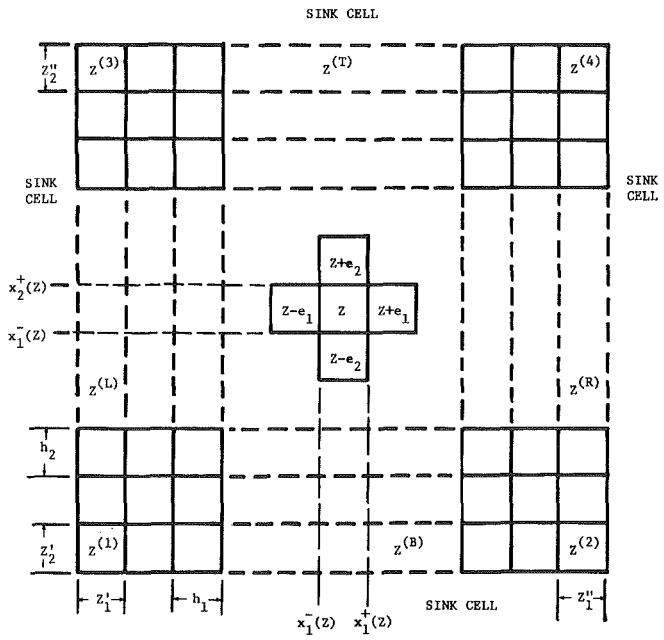


Fig. 1 Cell state space for two-dimensional systems

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-\frac{dt}{h_1 h_2}\int_{x_2^-}^{x_2^+} F_1(x_1^+,x_2)S(F_1(x_1^+,x_2);M(\mathbf{Z}+\mathbf{e}_1),M(\mathbf{Z}))\ dx_2, \tag{28a}$$

$$\frac{dt}{h_1 h_2}\int_{x_1^-}^{x_1^+} F_2(x_1,x_2^-)S(F_2(x_1,x_2^-);M(\mathbf{Z}),M(\mathbf{Z}-\mathbf{e}_2))\ dx_1, \tag{28b}$$

$$-\frac{dt}{h_1 h_2}\int_{x_1^-}^{x_1^+} F_2(x_1,x_2^+)S(F_2(x_1,x_2^+);M(\mathbf{Z}+\mathbf{e}_2),M(\mathbf{Z}))\ dx_1. \tag{28c}$$

The net gain of trajectories by the cell $\mathbf{Z}$ during dt is the sum of these four integrals which *can* be evaluated readily when $\mathbf{F}$ is given[1]. Expressed in terms of the cell probability function, the result can be put in the form:

$$\frac{d\zeta(\mathbf{Z})}{dt} = q_{\mathbf{Z},\mathbf{Z}-\mathbf{e}_1}(t)\,\zeta(\mathbf{Z}-\mathbf{e}_1)+q_{\mathbf{Z},\mathbf{Z}-\mathbf{e}_2}(t)\,\zeta(\mathbf{Z}-\mathbf{e}_2) + q_{\mathbf{Z},\mathbf{Z}}(t)\,\zeta(\mathbf{Z})+q_{\mathbf{Z},\mathbf{Z}+\mathbf{e}_1}(t)\,\zeta(\mathbf{Z}+\mathbf{e}_1) + q_{\mathbf{Z},\mathbf{Z}+\mathbf{e}_2}(t)\,\zeta(\mathbf{Z}+\mathbf{e}_2) \tag{29}$$

This is the dynamic equation for the cell probability function of an interior cell $\mathbf{Z}$.

Boundary Cells and Corner Cells. For a boundary cell bordering the left edge, but not the corner cells, the integrals of (28) remain unchanged, but the integral (27) is to be modified by replacing the switching function there by

$$S(F_1(x_1^-,x_2);M(\mathbf{Z}),0) \tag{30}$$

Similar modifications are to be applied to boundary cells bordering the right, bottom, and top edges.

For a corner cell, two of the four integrals need modification. For instance, if the cell is the left lower corner cell, then (28*a*) and (28*c*) remain unchanged but the integral (27) should be modified as in the case of a left edge cell and the integral (28*b*) is to be modified by using a new switching function

$$S(F_2(x_1,x_2^-);M(\mathbf{Z}),0). \tag{31}$$

Sink Cell. For the sink cell the rule of evolution is the following:

$$\frac{d\zeta(\mathbf{Z}_S)}{dt} = -\int_{x_2^-(\mathbf{Z}^{(1)})}^{x_2^+(\mathbf{Z}^{(1)})} F_1(x_1^-(\mathbf{Z}^{(1)}),x_2)\, S(F_1(x_1^-(\mathbf{Z}^{(1)}),x_2);\zeta(\mathbf{Z}^{(1)}),0)\ dx_2 - \int_{x_1^-(\mathbf{Z}^{(1)})}^{x_1^+(\mathbf{Z}^{(1)})} F_2(x_1,x_2^-(\mathbf{Z}^{(1)}))S(F_2(x_1,x_2^-(\mathbf{Z}^{(1)}));\zeta(\mathbf{Z}^{(1)}),0)\ dx_1$$
$$+ \text{contributions from other corner cells } \mathbf{Z}^{(2)},\mathbf{Z}^{(3)},\mathbf{Z}^{(4)}$$
$$- \sum_{\mathbf{Z}^{(L)}} \int_{x_2^-(\mathbf{Z}^{(L)})}^{x_2^+(\mathbf{Z}^{(L)})} F_1(x_1^-(\mathbf{Z}^{(L)}),x_2)\, S(F_1(x_1^-(\mathbf{Z}^{(L)}),x_2);\zeta(\mathbf{Z}^{(L)}),0)\ dx_2$$
$$+ \text{contributions from all the other edge cells } \mathbf{Z}^{(R)},\ \mathbf{Z}^{(B)},\ \mathbf{Z}^{(T)}. \tag{32}$$

where $\mathbf{Z}^{(1)}$, $\mathbf{Z}^{(2)}$, $\mathbf{Z}^{(3)}$, and $\mathbf{Z}^{(4)}$ stand for the left lower, right lower, left top, and right top corner cells, and $\mathbf{Z}^{(L)}$, $\mathbf{Z}^{(R)}$, $\mathbf{Z}^{(B)}$, and $\mathbf{Z}^{(T)}$ stand for all the left, right, bottom, and top edge cells, excluding the corner cells.

If we now put the sink cell and the $(Z_2''-Z_2'+1)(Z_1''-Z_1'+1)$ regular cells into a one-dimensional array and introduce a cell probability vector ζ with components $\zeta_i(t)$ indexed according to this array, then we are led again to an evolution equation for the cell probability vector

$$\frac{d\zeta}{dt} = \mathbf{Q}(t)\,\zeta \tag{33}$$

where $\mathbf{Q}(t)$ will again be called the cell probability flow-rate matrix.

Properties of the Rate Matrix Q. The rate matrix of (33) has the following properties:

1. The first column, associated with the sink cell, is a zero column.
2. Each row, except the first, can at most have five nonzero elements.
3. Each column can have at most five nonzero elements.
4. All off-diagonal elements are non-negative and all diagonal elements are nonpositive.
5. Each column has a zero sum.

4 Intensity Matrices and Stochastic Matrices

In this section we describe for the convenience of reference the properties of certain matrices that have special names in the mathematical literature.

Intensity Matrix. A matrix $\mathbf{A}$ is called an *intensity matrix* [4] if it has the following properties:

(*i*) All the off-diagonal elements are non-negative.

(*ii*) All the diagonal elements are nonpositive.

(*iii*) Each column has a zero sum. Because of (*i*) and (*ii*) this also implies

$$a_{jj} = -\sum_{\substack{i\\ i\neq j}} a_{ij} \tag{34}$$

Obviously, the rate matrix $\mathbf{Q}$ derived in Sections 2 and 3 is an intensity matrix. Further properties of an intensity matrix may be derived.

(*A*) If $\mathbf{E}$ denotes a column vector with all elements equal to unity, then because of (*iii*)

$$\mathbf{E}^T\mathbf{A} = 0 \cdot \mathbf{E}^T = \mathbf{0}. \tag{35}$$

Therefore, $\lambda=0$ is an eigenvalue of $\mathbf{A}$, and $\mathbf{E}$ is a left-eigenvector associated with the zero eigenvalue. Let the corresponding right-eigenvector be denoted by $\mathbf{u}$ so that

$$\mathbf{Au} = \mathbf{0}. \tag{36}$$

Of course, $\lambda=0$ could be a multiple eigenvalue and there can be several eigenvectors for $\lambda=0$.

(*B*) By (*i*), (*ii*), and (*iii*) it follows immediately that if $a_{jj}=0$ then the jth column of an intensity matrix is a zero column.

(*C*) It can also be shown that all the eigenvalues of an intensity matrix have nonpositive real parts.

Subintensity Matrix. For our purpose we call a matrix a *subintensity matrix* if it satisfies (*i*) and (*ii*) but with (*iii*) is replaced by

$$a_{jj} \leqq -\sum_{\substack{i\\ i\neq j}} a_{ij} \tag{37}$$

If, in our derivation leading to (21) and (33) the sink cells are deleted from the cell array for the cell probability vector ζ, then the rate matrices $\mathbf{Q}$ in those equations will be subintensity matrices in general.

Stochastic Matrix. A matrix $\mathbf{B}$ is called a *stochastic Matrix* [4–7] if

$$b_{ij} \geqq 0, \quad \text{and} \quad \sum_i b_{ij} = 1. \tag{38}$$

Substochastic Matrix. A matrix $\mathbf{B}$ is called a *substochastic matrix* if

[1] To evaluate the integrals one could use simplifying approximations that are consistent with the approximation caused by discretizing the state space with a cell structure.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$b_{ij} \geqq 0, \quad \text{and} \quad \sum_i b_{ij} \leqq 1. \tag{39}$$

5 General Solution and Global Behavior

When the system behavior is formulated in terms of the cell probability distribution, the evolution of the system is expressed in the form of (33). It is interesting to note that while the governing differential equation (3) or (23) may be highly nonlinear the *probability* evolution equation is *always linear*. One, of course, pays a penalty for this advantage by having an evolution equation of extremely high order. However, there is a consolation factor in that the rate matrix $\mathbf{Q}$ is usually quite sparse.

Let $\mathbf{P}(t)$ be the fundamental matrix of (33) with

$$\mathbf{P}(0) = \mathbf{I} \tag{40}$$

where $\mathbf{I}$ is a unit matrix. If one is given an initial cell probability distribution $\zeta(0)$, then the solution of (33) is given by

$$\zeta(t) = \mathbf{P}(t)\,\zeta(0) \tag{41}$$

Two particular types of solutions are of interest. One is for

$$\zeta(0) = \{0, \ldots, 0, \underset{(i)}{1}, 0, \ldots, 0\}^T \tag{42}$$

i.e., the ensemble of trajectories all start from cell i. The solution (41) in the form of the ith column of $\mathbf{P}(t)$ will tell us how the cell probability will be redistributed among the cells as time proceeds. The second problem is to have

$$\zeta(0) = \{0, \epsilon, \epsilon, \ldots, \epsilon\}^T \tag{43}$$

where we have placed the sink cell at the first position in the array and the initial probability distribution is uniform among all the regular cells, i.e.,

$$\epsilon = 1/(\text{number of regular cells}) \tag{44}$$

As time increases, the solution (41) will show us how this uniformly distributed "cloud" of probability will gradually concentrate toward the sink cell, a few absorbing cells, or a few groups of cells.

As for the fundamental matrix $\mathbf{P}(t)$ we have the following important result.

Theorem 1. The fundamental matrix $\mathbf{P}(t)$ of (33) and (40) where $\mathbf{Q}(t)$ is an intensity matrix is a stochastic matrix.

Proof. The proof is patterned after that given in [7] for the case where $\mathbf{Q}$ is a constant matrix. $\mathbf{P}(t)$ satisfies

$$\dot{P}(t) = \mathbf{Q}(t)\mathbf{P}(t) \quad \text{with} \quad \mathbf{P}(0) = \mathbf{I} \tag{45}$$

where we have used the dot to denote the time differentiation. Let $\boldsymbol{\Psi}$ be a diagonal matrix with

$$\psi_{ii} = \exp \int_0^t -q_{ii}(s)\,ds. \tag{46}$$

$\dot{\boldsymbol{\Psi}}$ is also a diagonal matrix. Moreover, it is equal to

$$\dot{\boldsymbol{\Psi}}(t) = -\boldsymbol{\Psi}(t)\mathbf{Q}_D(t) \tag{47}$$

where $\mathbf{Q}_D(t)$ denotes the diagonal part of $\mathbf{Q}(t)$. By (45) we have

$$\dot{\mathbf{P}}(t) = \mathbf{Q}_D(t)\mathbf{P}(t) + [\mathbf{Q}(t) - \mathbf{Q}_D(t)]\mathbf{P}(t) \tag{48}$$

Moving the first term on the right-hand side of the equation to the left-hand side, premultiplying by $\boldsymbol{\Psi}$, and making use of (47), one obtains

$$[\boldsymbol{\Psi}(t)\mathbf{P}(t)]^{\cdot} = \boldsymbol{\Psi}(t)[\mathbf{Q}(t) - \mathbf{Q}_D(t)]\mathbf{P}(t) \tag{49}$$

From (49) one readily find $\mathbf{P}(t)$ in the form

$$\mathbf{P}(t) = \boldsymbol{\Psi}^{-1}(t)\left\{\mathbf{I} + \int_0^t \boldsymbol{\Psi}(s)[\mathbf{Q}(s) - \mathbf{Q}_D(s)]\mathbf{P}(s)\,ds\right\} \tag{50}$$

By the theory of integral equations the solution $\mathbf{P}(t)$ of (50) can be put in the form

$$\mathbf{P}(t) = \sum_{n=0,1,\ldots} \mathbf{P}_n(t) \tag{51}$$

with

$$\mathbf{P}_0(t) = \boldsymbol{\Psi}^{-1}(t)$$

$$\mathbf{P}_1(t) = \boldsymbol{\Psi}^{-1}(t) \int_0^t \boldsymbol{\Psi}(s)[\mathbf{Q}(s) - \mathbf{Q}_D(s)]\mathbf{P}_0(s)\,ds$$

. . . .

$$\mathbf{P}_{n+1}(t) = \boldsymbol{\Psi}^{-1}(t) \int_0^t \boldsymbol{\Psi}(s)[\mathbf{Q}(s) - \mathbf{Q}_D(s)]\mathbf{P}_n(s)\,ds. \tag{52}$$

Since $\boldsymbol{\Psi}(t)$, $\boldsymbol{\Psi}^{-1}(t)$, and $\mathbf{Q}(t) - \mathbf{Q}_D(t)$ are all non-negative, all $P_n(t)$ are non-negative. Therefore, $\mathbf{P}(t)$ is non-negative. This meets the first requirement $p_{ij}(t) \geqq 0$ for $\mathbf{P}(t)$ being a stochastic matrix.

Premulitplying (45) by $\mathbf{E}^T$, one obtains

$$\mathbf{E}^T\dot{\mathbf{P}}(t) = \mathbf{E}^T\mathbf{Q}(t)\mathbf{P}(t) = \mathbf{0} \tag{53}$$

because $\mathbf{E}^T\mathbf{Q} = \mathbf{0}$. Hence,

$$\mathbf{E}^T\mathbf{P}(t) = \text{constant}. \tag{54}$$

But $\mathbf{E}^T\mathbf{P}(0) = \mathbf{E}^T$. Therefore,

$$\mathbf{E}^T\mathbf{P}(t) = \mathbf{E}^T. \tag{55}$$

This shows that the sum of each column of $\mathbf{P}(t)$ is unity. This completes the proof that $\mathbf{P}(t)$ is a stochastic matrix.

6 The Autonomous Systems

For autonomous systems $\mathbf{F}(\mathbf{x})$ is independent of t. An immediate consequence is that the cell probability flow-rate matrix $\mathbf{Q}$ is now a *constant* matrix. In that case the fundamental matrix $\mathbf{P}(t)$ discussed in Section 5 takes the form

$$\mathbf{P}(t) = \exp(\mathbf{Q}t) \tag{56}$$

While this simple expression contains, in principle, all the properties of the system we are looking for, it is not particularly useful until more details are known explicitly about $\mathbf{P}(t)$. To have a better appreciation about the global behavior of the system it is more instructive to study the structure of the matrix of $\mathbf{Q}$. Some of the discussions given in the following will follow the same line of thinking as that given in [3] for Markov chains. The reader might also wish to consult [4–6] for the mathematical theory of Markov chains.

Absorbing Cells. $\mathbf{Q}$ may have zero columns. Let us assume these columns are associated with cells $A_1, A_2, \ldots, A_\alpha$. A zero A_1 column means there is no flow of trajectories *leaving* cell A_1. Cell A_1 is, therefore, an absorbing cell. If, at the same time, A_1 row is not a zero row, there are cells from which the trajectories flow into cell A_1. In that case because of the inflow of trajectories into this cell, the cell probability ζ_{A_1} will increase monotonically until the sources are exhausted. Thus, if there are α number of zero columns in $\mathbf{Q}$, there are α number of absorbing cells. By our definition the sink cell introduced in this theory is always an absorbing cell.

Communicativeness. A useful qualitative property among the cells is communicativeness. When $i \neq j$, $q_{ij} \neq 0$ means that there is a positive flow rate of trajectories from cell j to cell i. In that case we say cell j *flows directly into* cell i. Let

$$\hat{\mathbf{Q}} = \mathbf{Q} - \mathbf{Q}_D \tag{57}$$

with elements $\hat{q}_{ij}$ ($\hat{q}_{jj} = 0$). Define, for $m, n \geqq 0$,

$$\hat{q}_{ij}^{(m+n)} = \sum_k \hat{q}_{ik}^{(m)}\hat{q}_{kj}^{(n)}, \quad \hat{\mathbf{Q}}^{(m+n)} = \hat{\mathbf{Q}}^{(m)}\hat{\mathbf{Q}}^{(n)} \tag{58}$$

with $\hat{q}_{ij}^{(0)} = \delta_{ij}$. If there exists a positive m such that $\hat{q}_{ij}^{(m)} > 0$, cell j is said to *flow to* cell i. The cells i and j are said to communicate if and only if cell i flows to cell j and cell j flows

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

to cell i. A cell that communicates with every cell it flows to is called *essential*. A cell that is not an absorbing cell nor an essential cell is called *inessential*.

Decomposition Into Groups. All the cells can be formed into groups. We have already seen the absorbing cells. We shall call each absorbing cell a group. If there are α absorbing cells, we have α groups. They will be labeled as $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_\alpha$.

Next, we divide all the essential cells into groups. The cells in one and the same group communicate but those belonging to different groups do not. These isolated groups will be called persistent groups. Let there be k such groups; they will be labeled as $\mathbf{B}_{\alpha+1}, \mathbf{B}_{\alpha+2}, \ldots, \mathbf{B}_{\alpha+k}$.

The inessential cells can also be formed into groups $\mathbf{B}_{\alpha+k+1}, \mathbf{B}_{\alpha+k+2}, \ldots$, and $\mathbf{B}_{\alpha+k+m}$ according to the rule that the system can flow from a cell in group $\mathbf{B}_{\alpha+k+h}$, $h=1,2,\ldots,m$,[2] to a cell in a group with a lower or equal subscript designation, but cannot flow to any group with a higher subscript designation. These groups will be called the transient groups. The decomposition of the whole cell space into groups also means the possibility of interchanging rows and interchanging columns in the rate matrix $\mathbf{Q}$ so that it will take the following *normal form*.

$$\mathbf{Q}=\begin{bmatrix}
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{S}_{1,\alpha+k+1} & \cdots & \mathbf{S}_{1,\alpha+k+m} \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{S}_{\alpha,\alpha+k+1} & \cdots & \mathbf{S}_{\alpha,\alpha+k+m} \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{Q}_{\alpha+1} & \cdots & \mathbf{0} & \mathbf{S}_{\alpha+1,\alpha+k+1} & \cdots & \mathbf{S}_{\alpha+1,\alpha+k+m} \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Q}_{\alpha+k} & \mathbf{S}_{\alpha+k,\alpha+k+1} & \cdots & \mathbf{S}_{\alpha+k,\alpha+k+m} \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}_{\alpha+k+1} & \cdots & \mathbf{S}_{\alpha+k+1,\alpha+k+m} \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_{\alpha+k+m}
\end{bmatrix} \tag{59}$$

where at the left there are α zero columns. $\mathbf{Q}_{\alpha+1}, \mathbf{Q}_{\alpha+2}, \ldots, \mathbf{Q}_{\alpha+k}$ are square matrices. Each of these is the cell probability flow-rate matrix for one of the persistent groups and, therefore, is an intensity matrix by itself. The matrices $\mathbf{R}_{\alpha+k+1}, \mathbf{R}_{\alpha+k+2}, \ldots$, and $\mathbf{R}_{\alpha+k+m}$ are associated with the transcient groups. Cells in each transient group communicate with each other but do not communicate with cells in other groups. Since there is a probability flow from a transient group to cells outside the group, the $\mathbf{R}$ matrices are not intensity matrices but subintensity matrices. The $\mathbf{S}$ matrices are in general rectangular. $\mathbf{S}_{ij}$, $i=1, 2, \ldots, \alpha+k$, and $j=\alpha+k+1, \alpha+k+2, \ldots, \alpha+k+m$, describe the manner by which the probability flows from the transient groups into the absorbing cells and the persistent groups. In a similar manner the $\mathbf{S}_{ij}$ matrices with $i=\alpha+k+1, \alpha+k+2, \ldots, \alpha+k+m-1$, and $j=\alpha+k+2, \alpha+k+3, \ldots, \alpha+k+m$ describe the manner of probability flow from the transient groups to transient groups of lower subscript designation. These $\mathbf{S}$ matrices will be called the transit matrices. When there is no advantage to have m distinct transient groups we shall lump all of them together and call it the transient group with a designation group $\mathbf{B}_{\alpha+k+1}$ and a subintensity matrix $\mathbf{R}_{\alpha+k+1}$. In that case the number of transit matrices will be simply $\alpha+k$ and they will be denoted by $\mathbf{S}_{1,\alpha+k+1}, \mathbf{S}_{2,\alpha+k+1}, \ldots$, and $\mathbf{S}_{\alpha+k,\alpha+k+1}$. From now on we shall also use the symbol $\mathbf{B}_j$ as a set notation to denote the set of all the cells in group $\mathbf{B}_j$.

[2] In the remainder of the paper we reserve α, k, and m to denote the numbers of the absorbing cells, persistent, and transient groups, respectively.

After the rate matrix $\mathbf{Q}$ has been put in normal form certain general overall behavior of the system can be readily discussed. For the convenience of the discussion to follow we assume that the cells all have been reindexed according to the order of positions they occupy in the normal form of (59).

1. Associated with each absorbing cell, say cell I, $1 \leq I \leq \alpha$, there is a *stationary* probability distribution for the system

$$\zeta=\{0, \ldots, 0, \underset{I\text{th}}{1}, 0, \ldots, 0\}^T \tag{60}$$

Also, from (33) and (59)

$$\dot{\zeta}_I=\sum_{j\in\mathbf{T}} q_{Ij}\zeta_j, \quad \mathbf{T}=\bigcup_{j=1}^{m}\mathbf{B}_{\alpha+k+j} \tag{61}$$

where $\mathbf{T}$ denotes the set of all transcient cells. It then follows that ζ_I is nondecreasing. In particular, if the system has its total probability 1 residing in cell I at any time, then $\zeta_1=1$ for all subsequent time.

2. Next, consider a persistent group, say group $\mathbf{B}_{\alpha+J}$, $1 \leq J \leq k$. Again, there is a *stationary* probability distribution for the system which is in the form

$$\zeta=\{0, \ldots, 0, \underset{\text{cells in } \mathbf{B}_{\alpha+J}}{p_1, p_2, \ldots, p_{n_J}}, 0, \ldots, 0\}^T \tag{62}$$

where n_J denotes the number of cells in $\mathbf{B}_{\alpha+J}$ and the p's give the stationary distribution among the cells in this group with

$$\sum_{i=1}^{n_J} p_i=1 \quad \text{and} \quad \mathbf{Q}_{\alpha+J}\begin{pmatrix} p_1 \\ \vdots \\ p_{n_J}\end{pmatrix}=0 \tag{63}$$

Obviously, $\{p_1, p_2, \ldots, p_{n_J}\}^T$ is the right-eigenvector associated with the zero eigenvalue of $\mathbf{Q}_{\alpha+J}$. Once $\mathbf{Q}_{\alpha+J}$ is known the stationary probability distribution can be calculated by solving (63).

Let us now denote by $\zeta(\mathbf{B}_{\alpha+J})$ the sum of the probabilities of all the cells in this group, i.e.,

$$\zeta(\mathbf{B}_{\alpha+J})=\sum_{i\in\mathbf{B}_{\alpha+J}}\zeta_i \tag{64}$$

Then by (33) and (59) we have

$$\dot{\zeta}(\mathbf{B}_{\alpha+J})=\sum_{i\in\mathbf{B}_{\alpha+J}}\sum_{l\in\mathbf{T}} q_{il}\zeta_l \geq 0 \tag{65}$$

Hence this group probability is nondecreasing. Again, if the whole ensemble of trajectories resides in this group at a certain time, then the ensemble remains inside this group thereafter.

3. In this subsection we consider a persistent group $\mathbf{B}_{\alpha+J}$ *by itself* with its intensity matrix $\mathbf{Q}_{\alpha+J}$. To simplify the notation we drop the group designation completely. We

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

simply use $\mathbf{Q}$ and ζ to denote the rate matrix and the cell probability vector. We note first that a persistent group cannot be further decomposed; hence it is usually called an irreducible set. The matrix $\mathbf{Q}$ being an intensity matrix of an irreducible set can be shown to have a simple zero eigenvalue and its eigenvector is $\mathbf{p}$ satisfying

$$\mathbf{Q}\,\mathbf{p}=\mathbf{0}, \quad \sum_i p_i = 1 \tag{66}$$

The time history of the cell probability vector is given by (41) with $\mathbf{P}(t)$ given by (56). If we examine $\zeta(t)$ at the instants τ, $2\tau, 3\tau, \ldots, n\tau, \ldots$ then we find

$$\zeta((n+1)\tau)=\mathbf{P}(\tau)\,\zeta(n\tau) \tag{67}$$

Therefore, $\zeta(n\tau)$ gives the evolution of an irreducible Markov chain [3-6] having $\mathbf{P}(\tau)$ as its transition probability matrix. Such a transition matrix is known to have a simple eigenvalue 1 with a unique eigenvector $\mathbf{p}^*$ which satisfies

$$\mathbf{P}(\tau)\,\mathbf{p}^*=\mathbf{p}^* \quad \sum_i p_i^*=1. \tag{68}$$

Since

$$\mathbf{P}(\tau)=\exp(\tau\,\mathbf{Q}), \tag{69}$$

one sees immediately that the eigenvector $\mathbf{p}$ of $\mathbf{Q}$ calculated from (66) is identical with the eigenvector $\mathbf{p}^*$ of $\mathbf{P}(\tau)$ calculated from (68). More importantly, this eigenvector is also the limiting probability distribution of the Markov chain (67), [3-6]. This result is independent of the size choice of τ. The preceding discussion indicates that the eigenvector $\mathbf{p}$ calculated for each of the persistent groups will play a most important role in the long-term behavior of the system.

4. Now we go back to the complete system and consider the transient groups. Here it is instructive to examine the structure of the fundamental matrix $\mathbf{P}(t)$ when $\mathbf{Q}$ is in its normal form. It is interesting to find that $\mathbf{P}(t)$ has the *same* general structure as $\mathbf{Q}$ and can be put in the form

$$\mathbf{P}(t) = \begin{bmatrix}
1 & \cdot & 0 & \mathbf{0} & \cdot & \mathbf{0} & \mathbf{T}_{1,\alpha+k+1} & \cdot & \mathbf{T}_{1,\alpha+k+m} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 1 & \mathbf{0} & \cdot & \mathbf{0} & \mathbf{T}_{\alpha,\alpha+k+1} & \cdot & \mathbf{T}_{\alpha,\alpha+k+m} \\
\mathbf{0} & \cdot & \mathbf{0} & \mathbf{P}_{\alpha+1} & \cdot & \mathbf{0} & \mathbf{T}_{\alpha+1,\alpha+k+1} & \cdot & \mathbf{T}_{\alpha+1,\alpha+k+m} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\mathbf{0} & \cdot & \mathbf{0} & \mathbf{0} & \cdot & \mathbf{P}_{\alpha+k} & \mathbf{T}_{\alpha+k,\alpha+k+1} & \cdot & \mathbf{T}_{\alpha+k,\alpha+k+m} \\
\mathbf{0} & \cdot & \mathbf{0} & \mathbf{0} & \cdot & \mathbf{0} & \mathbf{U}_{\alpha+k+1} & \cdot & \mathbf{T}_{\alpha+k+1,\alpha+k+m} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\mathbf{0} & \cdot & \mathbf{0} & \mathbf{0} & \cdot & \mathbf{0} & \mathbf{0} & \cdot & \mathbf{U}_{\alpha+k+m}
\end{bmatrix} \tag{70}$$

where the first α columns are zero columns except that there is a unit matrix of order $\alpha\times\alpha$ at the top. The next k columns have only block matrix elements at the diagonal positions with

$$\mathbf{P}_{\alpha+j}(t)=\exp(t\mathbf{Q}_{\alpha+j}), \quad j=1,2,\ldots,k. \tag{71}$$

The $\mathbf{U}$ matrices associated with the transient groups are given by

$$\mathbf{U}_{\alpha+k+j}(t)=\exp(t\mathbf{R}_{\alpha+k+j}), \quad j=1,2,\ldots,m. \tag{72}$$

The implications of the first $\alpha+k$ columns being populated by only 1's and $\mathbf{P}$'s at the diagonal positions have already been discussed in points 1-3.

With regard to the $\mathbf{U}$ matrices, we note that the $\mathbf{R}$ matrices are all subintensity matrices whose column sums are not all zero. It can be shown that for a matrix of this kind all the eigenvalues have negative real parts. Therefore,

$$\mathbf{U}_{\alpha+k+j}(t)\to\mathbf{0} \quad \text{as} \quad t\to\infty, \quad j=1,2,\ldots,m \tag{73}$$

As for the $\mathbf{T}$ matrices, one can evaluate them successively from the bottom up starting with the last transient group. Let us partition the vector $\zeta(t)$ into $\alpha+k+m$ parts, with ζ_j denoting the part associated with the jth group. Then for the last transient group $\mathbf{B}_{\alpha+k+m}$ we have by (41) and (70)

$$\zeta_{\alpha+k+m}(t)=\mathbf{U}_{\alpha+k+m}(t)\,\zeta_{\alpha+k+m}(0) \tag{74}$$

and

$$\zeta_{\alpha+k+m}(t)\to\mathbf{0} \quad \text{as} \quad t\to\infty. \tag{75}$$

Next, consider the group $\mathbf{B}_{\alpha+k+m-1}$. The solution is given by

$$\zeta_{\alpha+k+m-1}(t)=\mathbf{U}_{\alpha+k+m-1}(t)\left\{\zeta_{\alpha+k+m-1}(0)+\int_0^t \mathbf{U}^{-1}_{\alpha+k+m-1}(s)\,\mathbf{S}_{\alpha+k+m-1,\alpha+k+m}\,\mathbf{U}_{\alpha+k+m}(s)\,\zeta_{\alpha+k+m}(0)\,ds\right\} \tag{76}$$

Hence, after using (72), we have

$$\mathbf{T}_{\alpha+k+m-1,\alpha+k+m}(t)=\int_0^t \exp\{(t-s)\,\mathbf{R}_{\alpha+k+m-1}\}\,\mathbf{S}_{\alpha+k+m-1,\alpha+k+m}\exp(s\mathbf{R}_{\alpha+k+m})\,ds \tag{77}$$

Working backward from group to group we can express all the $\mathbf{T}_{j,l}$ in terms of $\mathbf{R}_j$, $\mathbf{Q}_j$, and $\mathbf{S}_{j,l}$ matrices of $\mathbf{Q}$. For all transient groups $\zeta_{\alpha+k+j}(t)\to\mathbf{0}$ as $t\to\infty$, $j=1,2,\ldots,m$. A stronger result is available for the last transient group. Let $\zeta(\mathbf{B}_{\alpha+k+m})$ be the sum of the probabilities of all the cells in group $\mathbf{B}_{\alpha+k+m}$. Then

$$\dot{\zeta}(B_{\alpha+k+m})=\sum_{i\in\mathbf{B}_{\alpha+k+m}}\sum_{j\in\mathbf{B}_{\alpha+k+m}} q_{ij}\zeta_j<0 \tag{78}$$

Hence the group probability of the last transient group not only goes to zero as $t\to\infty$ but also monotonically decreasing. This strong monotonic property may or may not be possessed by other transient groups. A transient group will have this property if there is no probability flow into that group.

From the preceding discussion, which is carried out in the spirit of the classical analyis of Markov chains, one might consider three tasks in the study of the global behavior of the system.

1. Determination of the absorbing cells, persistent groups, and the transient groups. When this task is accomplished the *qualitative* behavior of the system is mostly at hand, including the absorbing cells and/or persistent groups to which the probability from a transient cell will eventually flow.

2. Determination of the stationary probability distribution for each persistent group. This gives us the basis to calculate the eventual distribution of probability among the cells as $t\to\infty$ for a given initial probability distribution.

3. Determination of the long-term behavior of the $\mathbf{T}_{i,j}(t)$

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

block matrices with $i=1, 2, \ldots, \alpha+k$ and $j=\alpha+k+1, \alpha+k+2, \ldots, \alpha+k+m$ so that the time history of probability flow from each transient cell to the absorbing cells and persistent groups can be evaluated.

Here we recall that for the generalized cell mappings discussed in [3] similar tasks are necessary for the study of the global behavior. An algorithm for these tasks for the generalized cell mappings utilizing a very large number of cells has been developed and is reported in [8]. For the present theory, once the rate matrix $\mathbf{Q}$ is known, that algorithm with some slight modifications can be used to perform the task 1 and task 2 mentioned in the foregoing. For the task 3, a different algorithm is required.

7 Periodic Systems

When the dynamic system is nonautonomous the rate matrix $\mathbf{Q}$ depends on t. A discussion of this case is beyond the scope of this paper. We shall merely make a simple comment on periodic systems which will link the development here with the Markov chains discussed in [3]. Consider a periodic system such that $\mathbf{F}(\mathbf{x},t)$ of (23) is periodic in t of period τ_0. In that case $\mathbf{Q}(t)$ will also be periodic in t of same period. According to the Floquet theory the fundamental matrix $P(t)$ of (41) may be expressed as

$$\mathbf{P}(t)=\mathbf{W}(t)\ \mathbf{M}(t), \quad \mathbf{M}(t)=\exp\ (t\mathbf{R}) \tag{79}$$

where $\mathbf{W}(t)$ is a periodic matrix of period τ_0 with $\mathbf{W}(0)=\mathbf{I}$ and $\mathbf{R}$ is a constant matrix. Taking $t=\tau_0$, we have

$$\mathbf{P}(\tau_0)=\mathbf{W}(\tau_0)\ \mathbf{M}(\tau_0)=\mathbf{M}(\tau_0)=\exp\ (\tau_0\mathbf{R}) \tag{80}$$

If we now consider the cell probability vector at $t=n\tau_0$, $n=0, 1, 2, \ldots$, we find

$$\zeta((n+1)\tau_0)=\mathbf{M}(\tau_0)\ \zeta(n\tau_0). \tag{81}$$

This is a Markov chain with a stationary (constant) transition probability matrix $\mathbf{M}(\tau_0)$.

8. Concluding Remarks

In concluding, we may summarize the three basic features of the development presented here. First, one does not study individual trajectories but rather an ensemble of trajectories and uses a probability description for the state of the system. Second, the state space is structured into a cell space of large number of cells. This leads to a systems of *linear* ordinary differential equations with a sparse coefficient matrix. The third feature of the development is to use the theory of Markov processes.

One also notes here that a similar probabilistic theory for ensembles of trajectories can be developed by using the continuum state space. Such a theory leads to

$$\frac{\partial\zeta(\mathbf{x},t)}{\partial t}+\text{div}\{\zeta(\mathbf{x},t)\ \mathbf{F}(\mathbf{x},t)\}=0 \tag{82}$$

where *div* operation is taken with respect to the variables of the state space. This partial differential equation, while having a simple appearance, does not in general lend itself to simple analytical treatments. The present theory developed by using the cell state space may be regarded as a discretized version of (82).

This discretization entails, of course, approximation. A certain amount of diffusivity is built into the reformulated system by the discretization process. Thus, it is not expected for the theory to reproduce the results of the original systems in all the details. Perhaps, it could only be applied to certain kinds of systems to obtain certain types of information. There are also various possibilities of modifying the development. For instance, the cells do not need to be rectangular in shape. The sink cell could be chosen to be the part of the phase plane outside a closed curve which has the property that all trajectories cross it in the inward direction. Other modifications are possible. Therefore, this paper is to be taken merely as one in which we continue to explore the ramifications a cellularly structured state space may have. A great deal of further exploratory work remains to be done to test the viability and potential of this new approach to the global analysis of strongly nonlinear systems.

Acknowledgment

This material is based on work supported by the National Science Foundation under Grant No. MEA-8019274.

References

1 Hsu, C. S., "A Theory of Cell-to-Cell Mapping Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 931–939.

2 Hsu, C. S., and Guttalu, R. S., "Un Unravelling Algorithm for Global Analysis of Dynamical Systems: An Application of Cell-to-Cell Mappings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 940–948.

3 Hsu, C. S., "A Generalized Theory of Cell-to-Cell Mapping for Nonlinear Dynamical Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, 1981, pp. 634–642.

4 Isaacson, D. L., and Madsen, R. W., *Markov Chains: Theory and Applications*, Wiley, New York, 1976.

5 Chung, K. L., *Markov Chains With Stationary Transition Probabilities*, 2nd ed., Springer-Verlag, New York, 1967.

6 Romanovsky, V. I., *Discrete Markov Chains*, Wolters-Noordhoff, Groningen, The Netherlands, 1970.

7 Iosifescu, M., *Finite Markov Processes and Their Application*, Wiley, New York, 1980.

8 Hsu, C. S., Guttalu, R. S., and Zhu, W. H., "A Method of Analyzing Generalized Cell Mappings," ASME JOURNAL OF APPLIED MECHANICS, to appear.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Theory of Substructure Modal Synthesis

K. Kubomura
Rockwell International,
Space Transportation Systems,
12214 Lakewood Boulevard,
Downey, Calif. 90241

A theory is presented for representing the displacements of a substructure finite-element mathematical model with a reduced number of degrees of freedom. A first or second-order approximation is used for the substructure's modal coordinates associated with significantly larger or smaller eigenvalues than the system eigenvalues of excitation. The derived representations of the substructure displacements are capable of employing any type of substructure natural mode; free-free, cantilever or hybrid mode, and of retaining the dynamic behavior of any frequency range. It is shown that the present representations compute the system eigenvalues of interest with satisfactory accuracy, and it appears that the second-order approximation methods compute the system eigenvalues with greater accuracy than the first-order methods.

Introduction

Substructure synthesis is a method of matrix order reduction for structural dynamic analysis. A complete structural system is to be decomposed into several substructures. The motion of each substructure is represented by a set of substructure modes that accommodate the system eigenvalues and modes of interest. Free-free, constrained, or hybrid boundary conditions, i.e., some interface coordinates free and others fixed, may be used to determine the substructure modes independently. The original substructures are then represented by a reduced number of degrees of freedom, which are coupled together to result in a smaller problem size. Computationally, this represents a more desirable method of analysis than analyzing the complete structural system, as long as the reduced structure can adequately respond in the dynamic loading environment. Two basic questions have been directed: (1) how to select a set of substructure modes to obtain system eigenvalues and modes of interest with reasonable accuracy and computational cost; and (2) how to couple the substructures.

Hurty [12] described vibration of substructures using three modes: (1) cantilever modes, the lower natural modes of the substructure with all interface coordinates fixed; (2) rigid body modes; and (3) constraint modes. Lower cantilever modes are included in the transformation matrix intuitively, to retain the contents of the lower frequency modes. The displacement of the substructure due to interface coordinate displacements is accounted for by the introduction of rigid body and constraint modes. The reduced stiffness and mass matrices are labeled with modal coordinates and interface displacements, and coupled to adjacent substructures through continuity conditions of displacements, velocities, and accelerations across substructure boundaries. Craig and Bampton [2], realizing there is no need to treat rigid body modes separately from constraint modes, and proposed they both be regarded as constraint modes. Bamford [3] developed a computer program using cantilever modes and attachment modes for a constrained structure.

The vibration of substructures can also be described using free-free substructural modes, which are the natural modes of an unconstrained substructure. This approach has also been advocated by many authors [4–9] to retain lower frequency contents in reduced stiffness and mass matrices. Since interface displacements are not kept in the displacement transformation matrix directly as independent variables, special techniques are needed to couple the substructures. Benfield and Hruda [7], using stiffness and inertial loading at the substructure boundaries, attempted to account for the effect of adjacent substructures. Kuhar and Stahle [8] explain a method to use free-free modes of substructures that are interconnected by flexible links so that there are no common attachment points in the structural system. Hintz [9] extended attachment modes, initially developed for a constrained substructure, to free-free modes that are capable of producing the exact static response of a substructure to static loading at the interface points. The displacement transformation includes the boundary displacements before and after transformation, enabling the substructures to be coupled through the conditions of displacement continuity across the boundary. MacNeal [10] introduced a method that accounts for the contribution of the discarded higher modes called residual flexibility. The extension of MacNeal's work to a second-order approximation was discussed by Rubin [11].

Lower natural modes of the substructures are used to retain the lower frequency contents in the displacement transformation. The additional modes such as the constraint

Contributed by the Applied Mechanics Division and presented at the Winter Annual Meeting, Phoenix, Ariz., November 14–19, 1982 of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December, 1981; final revision, May, 1982. Paper No. 82-WA/APM-29.

Copies will be available until July, 1983.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

modes, the attachment modes, or the residual flexibilities, are introduced both to account for additional flexibility and to couple the substructures. In the present paper, two basic questions are addressed: (1) what the reason is for retaining certain component natural modes while approximating others; and (2) how to approximate certain modes. There is one clear criterion either to retain or to approximate modes; that is, the relative values of the substructure frequencies versus those of the system modes that may play a significant role in the dynamic environments of interest. Therefore, reduced representations are sought for the general coordinates associated with the substructure normal modes whose eigenvalues are either larger or smaller than the system eigenvalue of interest. First and second-order approximation methods for a component modes representation are presented in this paper.

First a method using free-free component modes is presented, then a method using cantilever modes is given. The use of hybrid modes is discussed, but only briefly, because the derivation of the method appears to be straightforward.

Free-Free Modes Representation

Second-Order General Representation. When a structural system is vibrating in a natural mode, the equation of motion of a substructure is, in the absence of damping forces,

$$\mathbf{M}\cdot\ddot{\mathbf{u}}+\mathbf{K}\cdot\mathbf{u}=\begin{Bmatrix}\mathbf{0}\\ \mathbf{f}_b\end{Bmatrix} \tag{1}$$

Here, $\mathbf{M}$ and $\mathbf{K}$ are the substructure mass and stiffness matrices, respectively, and $\mathbf{u}$ and $\ddot{\mathbf{u}}$ are the substructure displacement and acceleration vectors, respectively. Since the system structure is not subjected to any external forces, no forces are acting on the substructure except at its boundary points, which are denoted by a vector $\mathbf{f}_b$, subscript b indicating boundary degrees of freedom. For a substructure undergoing forced vibration of frequency Ω ($\Omega^2 = \Lambda$) the equation of motion, (1), becomes

$$(\mathbf{K}-\Lambda\mathbf{M})\cdot\mathbf{u}=\begin{Bmatrix}\mathbf{0}\\ {}^{\Lambda}\mathbf{f}_b\end{Bmatrix} \tag{2}$$

Displacements, $\mathbf{u}$, of the substructure are expressed as the linear combination of all free-free component natural modes, $\boldsymbol{\phi}$, normalized with respect to the mass matrix, with the general coordinate vector being $\mathbf{r}$, i.e.,

$$\mathbf{u}=\boldsymbol{\phi}\cdot\mathbf{r} \tag{3}$$

For the sake of convenience, four types of substructure modes will be classified; ${}^{R}\boldsymbol{\phi}$ are rigid body modes, ${}^{L}\boldsymbol{\phi}$ are lower modes whose associated eigenvalues, ${}^{L}\lambda$, are significantly lower than a system eigenvalue Λ of interest, ${}^{M}\boldsymbol{\phi}$ are modes that are retained since associated eigenvalues, ${}^{M}\lambda$, are comparable with Λ, and ${}^{H}\boldsymbol{\phi}$ are higher modes with eigenvalues, ${}^{H}\lambda$, being significantly larger than Λ. Equation (3) is rewritten with these four modes and their respective general coordinate vectors ${}^{R}\mathbf{r}$, ${}^{L}\mathbf{r}$, ${}^{M}\mathbf{r}$, and ${}^{H}\mathbf{r}$, i.e.,

$$\mathbf{u}={}^{R}\boldsymbol{\phi}\cdot{}^{R}\mathbf{r}+{}^{L}\boldsymbol{\phi}\cdot{}^{L}\mathbf{r}+{}^{M}\boldsymbol{\phi}\cdot{}^{M}\mathbf{r}+{}^{H}\boldsymbol{\phi}\cdot{}^{H}\mathbf{r} \tag{4}$$

To obtain the expression of a general coordinate r, the displacements of equation (3) are substituted into the equation of motion (2). Premultiplying it by an elastic mode vector ${}^{e}\boldsymbol{\phi}$, one obtains the following relation for the general coordinate of an elastic mode.

$$r=\frac{1}{\lambda-\Lambda}\,{}^{e}\boldsymbol{\phi}^{T}\cdot{}^{\Lambda}\mathbf{f}_b \tag{5}$$

This equation (5) is used to differentiate which general coordinates are approximated and which are retained. If there is a significant difference between the substructure eigenvalues, λ, and system eigenvalues, Λ, of equation (5), then the equation is approximated. Alternatively, if these values are comparable, the equation is not approximated and a new notation is used for the unapproximated general coordinate. For this purpose, equation (5) is rewritten using Taylor's expansion:

When $\lambda >> \Lambda$

$$r\equiv{}^{H}r=\frac{1}{{}^{H}\lambda}\left\{1+\frac{\Lambda}{{}^{H}\lambda}+0^2\left(\frac{\Lambda}{{}^{H}\lambda}\right)\right\}\cdot{}^{H}\boldsymbol{\phi}_b^{T}\cdot{}^{\Lambda}\mathbf{f}_b \tag{6}$$

and, when $\lambda << \Lambda$

$$r\equiv{}^{L}r=-\frac{1}{\Lambda}\left\{1+\frac{{}^{L}\lambda}{\Lambda}+0^2\left(\frac{{}^{L}\lambda}{\Lambda}\right)\right\}\cdot{}^{L}\boldsymbol{\phi}_b^{T}\cdot{}^{\Lambda}\mathbf{f}_b \tag{7}$$

When a substructure eigenvalue, λ, is comparable with system eigenvalue, Λ, no approximation is made:

$$r\equiv{}^{M}r \tag{8}$$

The general coordinates of higher and lower frequency substructure modes can be approximated by retaining the first or up to the second term in equations (6) and (7). With these approximations of the general coordinates, displacements of the substructure can be approximated to the extent that the substructure will respond adequately to the system excitation. Introducing the second-order approximations of equations (6) and (7) as well as equation (8) into the expression of the substructure displacements of equation (4), and summing all the contributions from the system modes of interest, the approximated component displacements are rewritten as

$$\begin{aligned}\mathbf{u}={}&{}^{R}\boldsymbol{\phi}\cdot{}^{R}\mathbf{r}-{}^{L}\boldsymbol{\phi}\cdot{}^{L}\boldsymbol{\phi}_b^{T}\cdot\mathbf{g}-{}^{L}\boldsymbol{\phi}\cdot{}^{L}\boldsymbol{\lambda}\cdot{}^{L}\boldsymbol{\phi}_b^{T}\cdot\mathbf{h}+{}^{M}\boldsymbol{\phi}\cdot{}^{M}\mathbf{r}\\ &+{}^{H}\boldsymbol{\phi}\cdot{}^{H}\boldsymbol{\lambda}^{-1}\cdot{}^{H}\boldsymbol{\phi}_b^{T}\cdot\mathbf{S}+{}^{H}\boldsymbol{\phi}\cdot{}^{H}\boldsymbol{\lambda}^{-2}\cdot{}^{H}\boldsymbol{\phi}_b^{T}\cdot\mathbf{t}\end{aligned} \tag{9}$$

where

$$\mathbf{g}=\sum_{\Lambda}\frac{{}^{\Lambda}\mathbf{f}_b}{\Lambda} \tag{10}$$

Nomenclature

- $\mathbf{u}$ = vector of displacements
- $\ddot{\mathbf{u}}$ = vector of accelerations
- $\mathbf{M}$ = mass matrix
- $\mathbf{K}$ = stiffness matrix
- $\mathbf{f}_b$ = vector of interface forces
- Λ = system eigenvalue
- $\boldsymbol{\Phi}$ = system mode
- $\boldsymbol{\phi}$ = substructure free-free mode
- λ = eigenvalues associated with $\boldsymbol{\phi}$
- $\bar{\boldsymbol{\phi}}$ = substructure cantilever mode
- $\bar{\lambda}$ = eigenvalues associated with $\bar{\boldsymbol{\phi}}$
- $(\ )_b$ = quantity pertinent to interface degrees of freedom
- $(\ )_i$ = quantity pertinent to internal degrees of freedom
- ${}^{\Lambda}(\ )$ = quantity pertinent to system eigenvalue
- $\bar{(\ )}$ = quantity pertinent to substructure cantilever mode
- ${}^{R}(\ )$ = quantity associated with rigid body modes
- ${}^{e}(\ )$ = quantity associated with elastic modes
- ${}^{L}(\ )$ = quantity associated with lower modes
- ${}^{M}(\ )$ = quantity associated with middle-range modes
- ${}^{H}(\ )$ = quantity associated with higher modes
- $\underset{\sim}{(\ )}$ = vector or matrix
- $(\ )^{T}$ = transpose of vector or matrix

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\mathbf{h} = \sum_{\Lambda} \frac{{}^{\Lambda}\mathbf{f}_b}{\Lambda^2} \tag{11}$$

$$\mathbf{S} = \sum_{\Lambda} {}^{\Lambda}\mathbf{f}_b \tag{12}$$

$$\mathbf{t} = \sum_{\Lambda} \Lambda {}^{\Lambda}\mathbf{f}_b \tag{13}$$

where λ is a diagonal matrix with the diagonal terms being eigenvalues, and λ^1 and λ^2 are also diagonal matrices with the diagonal terms being the inverses of the eigenvalues and the squares of the eigenvalues, respectively. Here, summations are taken over all the contributions from the system modes of interest. **S** of equation (12) is interpreted as the summation of the vectors of boundary forces acting on the substructure from adjacent substructures, and **t** of equation (13) is the summation of the vectors of accelerations resulting from the boundary forces s. **g** and **h** of equations (10) and (11) may be interpreted as double and quadruple integrations of the boundary forces with respect to time, respectively. Equation (9) will be discussed in more detail in the following sections.

Higher Modes Approximation. Substructure modes representation has been developed primarily to obtain reasonably accurate system lower modes and eigenvalues since in a great majority of linear dynamic problems, only lower system modes play an important role in the structure's response spectrum. Using a first-order approximation for higher modes of a substructure, that is, to retain the first, fourth, and fifth terms in equation (10), the expression of the substructure displacements will be

$$\mathbf{u} = {}^{R}\boldsymbol{\phi} \cdot {}^{R}\mathbf{r} + {}^{M}\boldsymbol{\phi} \cdot {}^{M}\mathbf{r} + {}^{H}\boldsymbol{\phi} \cdot {}^{H}\boldsymbol{\lambda}^{-1} \cdot {}^{H}\boldsymbol{\phi}_b^T \cdot \mathbf{S} \tag{14}$$

This expression of substructure displacements is equivalent to the one that is implicitly used in Rubin's work [11] and that is further discussed by Craig and Chang [12] in conjunction with the attachment modes expression. Expressing equation (8) in the following form:

$${}^{M}r = {}^{M}\beta + \frac{1}{{}^{M}\lambda} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{15}$$

where ${}^{M}\beta$ is an independent variable defined by

$$\beta = \frac{\Lambda}{{}^{M}\lambda({}^{M}\lambda - \Lambda)} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{16}$$

one may derive a first-order lower modes representation which is equivalent to the transformation used in Hintz's work [9]. Using the following relation with the attachment modes method of Hintz

$$\mathbf{K} \cdot {}^{e}\boldsymbol{\phi} \cdot {}^{e}\boldsymbol{\lambda}^{-1} \cdot {}^{e}\boldsymbol{\phi}^T + \mathbf{M} \cdot {}^{R}\boldsymbol{\phi} \cdot {}^{R}\boldsymbol{\phi}^T = \mathbf{I} \tag{17}$$

the transformation matrix for free-free component modes in NASTRAN Level 17.5 can be derived. A first-order approximation of a substructure's higher mode general coordinates by the following equation,

$${}^{H}r = \frac{1}{{}^{H}\lambda} {}^{H}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{18}$$

is equivalent to assuming that the inertial forces resulting from accelerations of the substructure displacements associated with higher modes are negligible compared to elastic forces.

The second-order approximation of higher modes is obtained by retaining the first, third, fourth, and fifth terms of equation (9). In the second-order approximation, when the general coordinates are represented by the following equation with a different independent variable ${}^{M}\delta$,

$${}^{M}r = {}^{M}\delta + \frac{1}{{}^{M}\lambda} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b + \frac{\Lambda}{{}^{M}\lambda^2} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{19}$$

instead of by equation (8), the resulting equation of coordinate transformation is an extension of the method of attachment modes to a higher order approximation. Using equation (17), one can express the aforementioned equation in terms of mass and stiffness matrices.

Lower Modes Approximation. In this section, equations of transformation retaining higher modes and approximating lower modes are derived and discussed. Discarding the third, fifth, and sixth terms of equation (9), one obtains a first-order approximation of the lower modes, i.e.,

$$\mathbf{u} = {}^{R}\boldsymbol{\phi} \cdot {}^{R}\mathbf{r} - {}^{L}\boldsymbol{\phi} \cdot {}^{L}\boldsymbol{\phi}_b \cdot \mathbf{g} + {}^{M}\boldsymbol{\phi} \cdot {}^{M}\mathbf{r} \tag{20}$$

The approximation of the lower modes by

$${}^{L}r = \frac{1}{\Lambda} {}^{L}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{21}$$

is equivalent to the assumption that elastic forces due to substructure displacements associated with lower modes are negligible when compared with inertia forces due to accelerations of these displacements. To obtain a more computationally convenient form of equation (20), the general coordinates ${}^{M}r$ may be expressed in a different form than equation (8) with a different independent variable ${}^{M}X$;

$${}^{M}r = {}^{M}X + \frac{1}{\Lambda} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{22}$$

As a result, one obtains an expression similar to that of attachment modes, from which, with the help of the relationship

$$\mathbf{M} \cdot \boldsymbol{\phi} \cdot \boldsymbol{\phi}^T = \mathbf{I} \tag{23}$$

the internally partitioned displacement vector $\mathbf{u}_i$, can be expressed in terms of general coordinates associated with the rigid body modes, higher modes, and the boundary displacements $\mathbf{u}_b$;

$$\begin{aligned} \mathbf{u}_i &= ({}^{R}\boldsymbol{\phi}_i + \mathbf{M}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot {}^{R}\boldsymbol{\phi}_b) \cdot {}^{R}\delta \\ &+ ({}^{M}\boldsymbol{\phi}_i + \mathbf{M}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot {}^{M}\boldsymbol{\phi}_b) \cdot {}^{M}\beta + \mathbf{M}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot \mathbf{u}_b \end{aligned} \tag{24}$$

A second-order approximation of lower modes is derived by retaining the first four terms of equation (9). Using a new independent variable ${}^{M}\theta$,

$${}^{M}r = {}^{M}\theta + \frac{1}{{}^{M}\lambda} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b + \frac{\Lambda}{{}^{M}\lambda^2} {}^{M}\boldsymbol{\phi}_b^T \cdot {}^{\Lambda}\mathbf{f}_b \tag{25}$$

a computationally more economical expression of the transformation is obtained by replacing terms containing lower modes with partitioned stiffness and mass matrices, realizing the following relation:

$${}^{e}\boldsymbol{\phi}^T \cdot \mathbf{K} \cdot {}^{e}\boldsymbol{\phi} = {}^{e}\boldsymbol{\lambda} \tag{26}$$

By retaining the first, second, fourth, and fifth terms of equation (9), a first-order general expression of coordinate transformation is obtained,

$$u = {}^{R}\boldsymbol{\phi} \cdot {}^{R}\mathbf{r} - {}^{L}\boldsymbol{\phi} \cdot {}^{L}\boldsymbol{\phi}_b \cdot \mathbf{g} + {}^{M}\boldsymbol{\phi} \cdot {}^{M}\mathbf{r} + {}^{H}\boldsymbol{\phi} \cdot {}^{H}\boldsymbol{\lambda}^{-1} \cdot {}^{H}\boldsymbol{\phi}_b^T \cdot S \tag{27}$$

This expression of substructure displacements is equivalent to the one used by Klosterman and McClelland [13]. It may not be computationally economical to calculate all the modes and eigenvalues needed in order to use the transformation of equation (27). Instead, modifications of the transformation equations (14) or (20) may be used by replacing ${}^{M}\boldsymbol{\phi}$ and ${}^{M}\mathbf{r}$ in these equations by either $[-{}^{L}\boldsymbol{\phi} \cdot {}^{L}\boldsymbol{\phi}_b^T, {}^{M}\boldsymbol{\phi}]$ and $[\mathbf{g}, {}^{M}\mathbf{r}]$, or $[{}^{M}\boldsymbol{\phi}, {}^{H}\boldsymbol{\phi} \cdot {}^{H}\boldsymbol{\lambda}^{-1} \cdot {}^{H}\boldsymbol{\phi}_b^T]$ and $[{}^{M}\mathbf{r}, \mathbf{S}]$, depending on the way the modes of the component structure are calculated (higher or lower modes).

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Cantilever Modes Representation

Second-Order General Representation. An approximation for the displacements of a substructure has been done using the free-free modes representation of the substructure in the preceding section. In this section, the approximation of internal displacements $\mathbf{u}_i$ is discussed. Displacements $\mathbf{u}_i$ at internal points are expressed as a linear combination of all natural modes of the substructure with all interface degrees of freedom constrained (cantilever modes), i.e.,

$$\mathbf{u}_i = \bar{\boldsymbol{\phi}} \cdot \bar{\mathbf{r}} \tag{28}$$

where $\bar{\boldsymbol{\phi}}$ is the cantilever modes matrix (normalized with respect to the partitioned mass matrix) and $\bar{\mathbf{r}}$ is the modal coordinate vector associated with $\bar{\boldsymbol{\phi}}$. The same notation that was used in deriving the free-free modes representations is used here except that the quantities pertinent to the cantilever modes are overscored, such as $\bar{\boldsymbol{\phi}}, \bar{\mathbf{r}}, \bar{\Lambda}$, etc.

The equation of motion of the substructure is described by equation (2), which can be partitioned as follows:

$$\begin{bmatrix} \mathbf{K}_{ii} - \Lambda \mathbf{M}_{ii} & \mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib} \\ \mathbf{K}_{bi} - \Lambda \mathbf{M}_{bi} & \mathbf{K}_{bb} - \Lambda \mathbf{M}_{bb} \end{bmatrix} \cdot \begin{Bmatrix} \mathbf{u}_i \\ \mathbf{u}_b \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ {}^{\Lambda}\mathbf{f}_b \end{Bmatrix} \tag{29}$$

The upper half of equation (29) is rewritten with the modification:

$$(\mathbf{K}_{ii} - \Lambda \mathbf{M}_{ii}) \cdot \mathbf{u}_i = -(\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{30}$$

where ${}^{\Lambda}\boldsymbol{\Phi}_b$ is the boundary partitioned system mode vector of excitation and ${}^{\Lambda}X$ is the general coordinate associated with the excited system mode. Introducing the expression of internal displacements of equation (28) into the partitioned equation of motion of the substructure (30), the general coordinates, $\bar{r}$, associated with cantilever modes are given by

$$\bar{r} = -\frac{1}{\bar{\lambda} - \Lambda} \bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{31}$$

General coordinates, $\bar{r}$, of the substructure modes with either larger or smaller eigenvalues relative to system eigenvalues of interest are approximated using the same method used for equations (6), (7), and (8); for $\bar{\lambda} >> \Lambda$

$$\bar{r} \equiv {}^{H}\bar{r} \doteq -\frac{1}{{}^{H}\bar{\lambda}} \left(1 + \frac{\Lambda}{{}^{H}\bar{\lambda}}\right) \bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{32}$$

for $\bar{\lambda} << \Lambda$

$$\bar{r} \equiv {}^{L}\bar{r} \doteq \frac{1}{\Lambda} \left(1 + \frac{{}^{L}\bar{\lambda}}{\Lambda}\right) {}^{L}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{33}$$

and for $\bar{\lambda} \sim \Lambda$

$$\bar{r} \equiv {}^{M}\bar{r} \tag{34}$$

Here, equation (34) indicates that the general coordinates associated with cantilever modes whose eigenvalues are comparable with system eigenvalues of interest are not approximated. Introducing second-order approximations for general coordinates of equations (32) and (33), and equation (34) into the expression of internal displacements of equation (28), and summing all of the contribution from the system modes of interest, the second-order approximation of internal displacements using cantilever modes is given by

$$\begin{aligned} \mathbf{u}_i = {} & {}^{L}\bar{\boldsymbol{\phi}} \cdot {}^{L}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b - \mathbf{M}_{ib} \cdot \mathbf{u}_b{}^{D}) \\ & + {}^{L}\bar{\boldsymbol{\phi}} \cdot {}^{L}\bar{\lambda} \cdot {}^{L}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b{}^{B} - \mathbf{M}_{ib} \cdot \mathbf{u}_b{}^{A}) + \bar{\boldsymbol{\phi}} \cdot \bar{r} \\ & - {}^{H}\bar{\boldsymbol{\phi}} \cdot {}^{H}\bar{\lambda}^{-1} \cdot {}^{H}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b^{C} - \mathbf{M}_{ib} \cdot \mathbf{u}_b{}^{D}) \\ & - \bar{\boldsymbol{\phi}} \cdot {}^{H}\bar{\lambda}^{-2} \cdot {}^{H}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b^{D} - \mathbf{M}_{ib} \cdot \mathbf{u}_b{}^{E}) \end{aligned} \tag{35}$$

where

$$\mathbf{u}_b^{A} = \sum_{\Lambda} {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{36}$$

$$\mathbf{u}_b^{B} = \sum_{\Lambda} \frac{1}{\Lambda} \boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{37}$$

$$\mathbf{u}_b^{C} = \sum_{\Lambda} \frac{1}{\Lambda^2} \, {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{38}$$

$$\mathbf{u}_b^{D} = \sum_{\Lambda} \Lambda \, {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{39}$$

$$\mathbf{u}_b^{E} = \sum_{\Lambda} \Lambda^2 \, {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{40}$$

It should be noted that $\mathbf{u}_b$ of equation (35) includes any contributions from rigid-body system modes.

Higher Modes Approximation. A method of substructure mode synthesis using cantilever modes was derived intuitively by Hurty and then extended by Craig-Bampton. This method can be derived as a special case of the general second-order cantilever modes approximation of equation (35). Discarding the first two terms and the last term of the right-hand side of equation (35), the first-order higher cantilever modes approximations becomes

$$\mathbf{u}_i = {}^{M}\bar{\boldsymbol{\phi}} \cdot {}^{M}\bar{r} - {}^{H}\bar{\boldsymbol{\phi}} \cdot {}^{H}\bar{\lambda}^{-1} \cdot {}^{H}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b - \mathbf{M}_{ib} \cdot \mathbf{u}_b^{D}) \tag{41}$$

While deriving equation (41), the notation of equation (38) was used. Using the following relation

$${}^{M}\bar{r} = {}^{M}\bar{\beta} + \frac{1}{{}^{H}\bar{\lambda}} \, {}^{H}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X \tag{42}$$

where

$$M_{\bar{\beta}} = \frac{\Lambda}{{}^{M}\bar{\lambda}({}^{M}\bar{\lambda} - \Lambda)} \, {}^{M}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X$$

equation (41) becomes

$$\mathbf{u}_i = {}^{M}\bar{\boldsymbol{\phi}} \, {}^{M}\bar{\beta} - \bar{\boldsymbol{\phi}} \cdot \bar{\lambda}^{-1} \cdot \boldsymbol{\Phi}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b - \mathbf{M}_{ib} \cdot \mathbf{u}_b^{D}) \tag{43}$$

Noting that $\mathbf{K}_{ii}^{-1} = \bar{\boldsymbol{\phi}} \cdot \bar{\lambda}^{-1} \cdot \bar{\boldsymbol{\phi}}^T$, the transformation of equation (33) becomes

$$\mathbf{u}_i = {}^{M}\bar{\boldsymbol{\phi}} \cdot {}^{M}\bar{\beta} - \mathbf{K}_{ii}^{-1} \cdot \mathbf{K}_{ib} \cdot \mathbf{u}_b + \mathbf{K}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot \mathbf{u}_b^{D} \tag{44}$$

It can be recognized that equation (44) without the third term is the same as the equation used by Craig-Bampton.

A similar expression as equation (44) for the second-order higher modes approximation is given by

$$\begin{aligned} \mathbf{u}_i = {} & {}^{M}\bar{\boldsymbol{\phi}} \cdot {}^{M}\bar{\mathbf{r}} - \mathbf{K}_{ii}^{-1} \cdot \mathbf{K}_{ib} \cdot \mathbf{u}_b \\ & + (\mathbf{K}_{ii}^{-1} \cdot \mathbf{M}_{ib} - \mathbf{K}_{ii}^{-1} \cdot \mathbf{M}_{ii} \cdot \mathbf{K}_{ii}^{-1} \cdot \mathbf{K}_{ib}) \cdot \mathbf{u}_b^{D} \\ & + \mathbf{K}_{ii}^{-1} \cdot \mathbf{M}_{ii} \cdot \mathbf{K}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot \mathbf{u}_b^{E} \end{aligned} \tag{45}$$

Lower Cantilever Modes Approximation. With a similar process as the one used for approximating higher modes, one can obtain first and second-order approximations of lower modes. Retaining the first and third terms of equation (35), the coordinate transformation of the first-order approximation is

$$\mathbf{u}_i = {}^{L}\bar{\boldsymbol{\phi}} \cdot {}^{L}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} \cdot \mathbf{u}_b - \mathbf{M}_{ib} \cdot \mathbf{u}_b^{D}) + {}^{M}\bar{\boldsymbol{\phi}} \cdot {}^{M}\bar{\mathbf{r}} \tag{46}$$

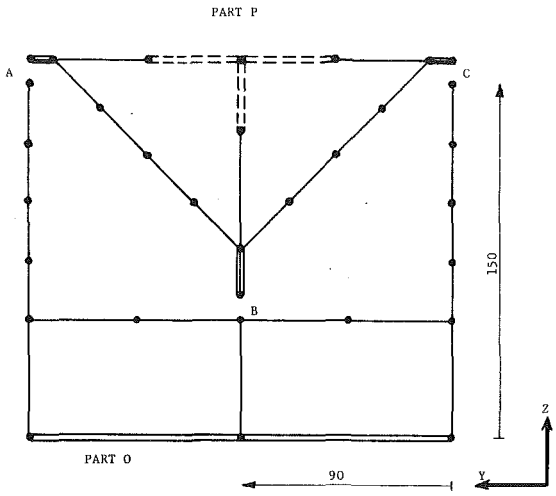
The first term of equation (46) represents the contribution from approximated lower modes and might be called "residual induced acceleration."

The attachment modes are calculated by adding retained flexibility to residual flexibility. By a similar procedure, rewriting equation (34) in the form (using a independent variable ${}^{M}\bar{\delta}$)

$${}^{M}\bar{r} = {}^{M}\bar{\delta} + \frac{1}{\Lambda} \, {}^{M}\bar{\boldsymbol{\phi}}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^{\Lambda}\boldsymbol{\Phi}_b \, {}^{\Lambda}X$$

the following expression can be obtained:

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



MATERIAL AND GEOMETRICAL PROPERTIES

ELEMENT	YOUNG'S MODULUS	POISSON'S RATIO	MASS DENSITY	CROSS SECTIONAL AREA	MOMENT OF INERTIA
	10 E+7	.3	.000259	8.6	32.9
	10 E+7	.3	.0259	8.6	32.9
	10 E+7	.3	.000259	28.3	63.6
	10 E+7	.3	.00259	28.3	63.6

Fig. 1 Example structure

$$\mathbf{u}_i = {}^M\bar{\phi} \cdot {}^M\bar{\delta} + \mathbf{M}_{ii}^{-1} \cdot \mathbf{K}_{ib} \cdot \mathbf{u}_b - \mathbf{M}_{ii}^{-1} \cdot \mathbf{M}_{ib} \cdot \mathbf{u}_b^D \tag{47}$$

Here, the relation

$$\mathbf{M}_{ii} \cdot \bar{\phi} \cdot \bar{\phi}^T = \mathbf{I} \tag{48}$$

was used. The coordinate transformation of the second-order lower modes approximation in terms of partitioned stiffness and mass matrices can also be obtained from equation (35), with the following expression of general coordinates using a different independent variable ${}^M\bar{t}$,

$$\begin{aligned} {}^M\bar{r} \;=\; & {}^M\bar{t} + \frac{1}{\Lambda}\, {}^M\bar{\phi}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^\Lambda\Phi_b \, {}^\Lambda X \\ & - \frac{{}^M\bar{\lambda}}{\Lambda^2}\, {}^M\bar{\phi}^T \cdot (\mathbf{K}_{ib} - \Lambda \mathbf{M}_{ib}) \cdot {}^\Lambda\bar{\phi}_b \, {}^\Lambda X \end{aligned} \tag{49}$$

The first-order approximation of lower cantilever modes is equivalent to the assumption that internal elastic forces due to the excitation of higher cantilever modes are negligible compared with corresponding inertial forces.

Hybrid mode representations of a substructure's displacements can also be derived with the first or second-order approximation of substructure hybrid mode general coordinates. The general coordinates that are approximated are those associated with the substructure hybrid modes of significantly higher or lower frequencies than the system frequencies of excitation. The substructure displacements at the unconstrained degrees of freedom are expressed with these approximated general coordinates, resulting in expressions with reduced degrees of freedom. From these expressions, the interface displacements are solved and are used as new independent variables in the new expression of coordinate transformations.

As have been discussed in this paper, there exist for every substructure two types of modes: the natural and interface modes. The interface modes are modes that compensate for the influence of the discarded natural modes and are used to couple substructure stiffness and mass matrices (e.g., attachment modes, residual modes, constraint modes). The natural modes, coupled with the interface modes, span a subspace in which the representation of stiffness and mass is

Table 1 Free-free modes approximation (approximated eigenvalues)

MODE	NO APPROXIMATION	RETAINED MODES	LOWER MODES	MIDDLE MODES	HIGHER MODES
4	7.5544 E+3		7.5544 E+3		
5	1.3429 E+4		1.3429 E+4		
6	2.5084 E+4		2.5084 E+4		
7	5.3542 E+4		5.3542 E+4		
8	8.1706 E+4		8.1706 E+4		
9	1.1087 E+5		1.1087 E+5		
10	2.0870 E+5		2.0872 E+5		
11	2.4454 E+5		2.4454 E+5		
12	2.5317 E+5		2.5317 E+5		
13	3.3864 E+5		3.3866 E+5		
14	3.4212 E+5		3.4213 E+5	1.1870 E+4	
15	3.9339 E+5		3.9352 E+5	4.9993 E+4	
16	5.4582 E+5		5.4589 E+5	1.2241 E+5	
17	6.1587 E+5		6.1657 E+5	1.5423 E+5	
18	8.3438 E+5		8.3443 E+5	2.1289 E+5	
19	9.8139 E+5		9.8340 E+5	2.5292 E+5	
20	1.1232 E+6		1.1233 E+6	3.0709 E+5	
21	1.1781 E+6		1.1793 E+6	5.4796 E+5	
22	1.3043 E+6		1.3491 E+6	7.3507 E+5	
23	2.0700 E+6		2.6807 E+6	8.4083 E+5	
24	2.6398 E+6		2.8500 E+6	1.1241 E+6	
25	2.6833 E+6		2.9467 E+6	1.3826 E+6	
26	2.8494 E+6		3.3807 E+6	1.4462 E+6	
27	2.9434 E+6		3.9462 E+6	2.6709 E+6	
28	2.0970 E+6		6.0172 E+6	2.8536 E+6	
29	3.2705 E+6		6.0823 E+6	2.9314 E+6	
30	3.5309 E+6		8.8005 E+6	3.5131 E+6	
31	5.0828 E+6		1.0606 E+7	3.6934 E+6	
32	5.7278 E+6		1.2226 E+7	5.7252 E+6	
33	5.7661 E+6		1.3061 E+7	5.7273 E+6	1.5911 E+4
34	7.2262 E+6		1.7452 E+7	7.2248 E+6	1.5443 E+5
35	8.2386 E+6		1.7558 E+7	8.1721 E+6	1.8366 E+5
36	1.0198 E+7		2.6195 E+7	1.0198 E+7	2.5960 E+5
37	1.0481 E+7		2.6815 E+7	1.0475 E+7	3.4671 E+5
38	1.2928 E+7		5.6056 E+7	1.2928 E+7	4.6280 E+5
39	1.3311 E+7		7.3202 E+7	1.3309 E+7	4.6832 E+5
40	1.5328 E+7		7.7520 E+7	1.5330 E+7	8.7739 E+5
41	1.7398 E+7		8.1038 E+7	1.7396 E+7	1.1284 E+6
42	1.8254 E+7		1.8189 E+8	1.8245 E+7	1.7458 E+6
43	1.9391 E+7		1.8495 E+8	1.9393 E+7	2.7278 E+6
44	2.8946 E+7		3.0809 E+8	2.8955 E+7	2.8633 E+6
45	2.9429 E+7		3.0984 E+8	2.9440 E+7	3.0369 E+6
46	3.9125 E+7			3.9445 E+7	5.1311 E+6
47	4.5515 E+7			4.5710 E+7	6.9170 E+6
48	5.2157 E+7			5.6061 E+7	7.1453 E+6
49	5.3341 E+7			5.8730 E+7	1.1604 E+7
50	5.6068 E+7			7.1377 E+7	1.2964 E+7
51	6.7290 E+7			7.8972 E+7	1.5830 E+7
52	7.8922 E+7			8.2759 E+7	2.7236 E+7
53	8.4168 E+7			1.0771 E+8	2.7511 E+7
54	8.5986 E+7			1.4619 E+8	3.2676 E+7
55	1.1893 E+8			1.5753 E+8	3.5317 E+7
56	1.2829 E+8			2.0564 E+8	3.5411 E+7
57	1.3956 E+8			2.2124 E+8	5.6065 E+7
58	1.5450 E+8			2.8147 E+8	7.6825 E+7
59	1.8231 E+8			3.1530 E+8	8.6015 E+7
60	1.8360 E+8			3.3046 E+8	8.6260 E+7
61	2.0467 E+8			5.8018 E+8	1.8105 E+8
62	2.3434 E+8				1.8119 E+8
63	2.4810 E+8				2.2227 E+8
64	2.9625 E+8				2.9582 E+8
65	2.9648 E+8				2.9595 E+8
66	3.5567 E+8				3.5562 E+8
67	3.5898 E+8				3.5879 E+8
68	4.6278 E+8				4.6763 E+8
69	4.6533 E+8				4.6509 E+8
70	5.4980 E+8				5.4737 E+8
71	1.0844 E+9				1.0844 E+9
72	1.0844 E+9				1.0844 E+9
73	8.0293 E+9				8.0293 E+9
74	8.0418 E+9				8.0418 E+9

RANGE OF RETAINED LOWER MODES: 1.4463 E+6
RANGE OF RETAINED MIDDLE MODES: 6.3146 E+6
RANGE OF RETAINED MIDDLE MODES: 4.9401 E+7
RANGE OF RETAINED HIGHER MODES: 3.4564 E+8

capable of retaining the substructure's necessary dynamic behavior. Therefore, a consistent choice of these two types of modes should be made for constructing the coordinate transformation to span an adequate subspace; for example, a choice of free-free natural modes should come with attachment or residual modes. Some coordinate transformations derived in this paper do not explicitly include the interface displacements, therefore, for the purpose of coupling the substructures, it must be converted to one that includes the interface displacements. This will be done by partitioning the coordinate transformation to the interface degrees of freedom and by solving the interface variables (e.g., $\mathbf{g}$ of equation (20)), $\mathbf{S}$ of equations (14) and (27)) in terms of the interface displacements $\mathbf{u}_b$. The converted transformation is applied either directly to the functional from which the system equation of motion is derived as an extremum condition, or to the substructure stiffness and mass matrices from which the system equation of motion is derived by coupling.

Example

A beam structure shown in Fig. 1 is chosen as an example problem so that the substructure modal representations developed here could be verified numerically. The structure

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 2 First and second-order approximation of higher modes (approximated eigenvalues)

NO. OF MODE	NO APPROXIMATION	FREE-FREE MODES APPROXIMATION		CANTILEVER MODES APPROXIMATION			
		1ST ORDER (RUBIN)	2ND ORDER	MODIFIED 1ST ORDER CRAIG-BAMPTON	1ST ORDER	MODIFIED 2ND ORDER	2ND ORDER
4	7.5544 E+3	7.5544 E+3	7.5544 E+3	7.5544 E+3	7.5544 E+3	7.5544 E+3	7.5544 E+3
5	1.3429 E+4	1.3429 E+4	1.3429 E+4	1.3429 E+4	1.3429 E+4	1.3429 E+4	1.3429 E+4
6	2.5084 E+4	2.5084 E+4	2.5084 E+4	2.5085 E+4	2.5084 E+4	2.5084 E+4	2.5084 E+4
7	5.3542 E+4	5.3542 E+4	5.3542 E+4	5.3554 E+4	5.3550 E+4	5.3542 E+4	5.3542 E+4
8	8.1706 E+4	8.1706 E+4	8.1706 E+4	8.1711 E+4	8.1709 E+4	8.1706 E+4	8.1706 E+4
9	1.1087 E+5	1.1087 E+5	1.1087 E+5	1.1090 E+5	1.1089 E+5	1.1087 E+5	1.1087 E+5
10	2.0870 E+5	2.0871 E+5	2.0870 E+5	2.0875 E+5	2.0874 E+5	2.0870 E+5	2.0870 E+5
11	2.4454 E+5	2.4454 E+5	2.4454 E+5	2.4457 E+5	2.4457 E+5	2.4454 E+5	2.4454 E+5
12	2.5317 E+5	2.5317 E+5	2.5317 E+5	2.5318 E+5	2.5317 E+5	2.5317 E+5	2.5317 E+5
13	3.3864 E+5	3.3866 E+5	3.3864 E+5	3.3868 E+5	3.3867 E+5	3.3864 E+5	3.3864 E+5
14	3.4212 E+5	3.4213 E+5	3.4212 E+5	3.4215 E+5	3.4213 E+5	3.4212 E+5	3.4212 E+5
15	3.9339 E+5	3.9353 E+5	3.9339 E+5	3.9368 E+5	3.9361 E+5	3.9339 E+5	3.9339 E+5
16	5.4582 E+5	5.4589 E+5	5.4582 E+5	5.4642 E+5	5.4596 E+5	5.4582 E+5	5.4582 E+5
17	6.1587 E+5	6.1657 E+5	6.1587 E+5	6.1681 E+5	6.1664 E+5	6.1588 E+5	6.1587 E+5
18	8.3438 E+5	8.3443 E+5	8.3438 E+5	8.3513 E+5	8.3501 E+5	8.3439 E+5	8.3438 E+5
19	9.8139 E+5	9.8340 E+5	8.8143 E+5	9.8982 E+5	8.8922 E+5	9.8154 E+5	9.8139 E+5
20	1.1232 E+6	1.1231 E+6	1.1232 E+6	1.1234 E+6	1.1233 E+6	1.1232 E+6	1.1232 E+6
21	1.1781 E+6	1.1793 E+6	1.1781 E+6	1.3083 E+6	1.3079 E+6	1.1836 E+6	1.1781 E+6
22	1.3043 E+6	1.3491 E+6	1.3049 E+6	2.5692 E+6	2.4768 E+6	1.3045 E+6	1.3043 E+6
23	2.0700 E+6	2.6807 E+6	2.3754 E+6	2.7414 E+6	2.7079 E+6	2.2178 E+6	2.0632 E+6
24	2.6398 E+6	2.8500 E+6	2.6746 E+6	2.8479 E+6	2.8454 E+6	2.6748 E+6	2.6675 E+6
25	2.6833 E+6	2.9467 E+6	2.8572 E+6	3.3342 E+6	3.1442 E+6	2.8544 E+6	2.7596 E+6
26	2.8494 E+6	3.3807 E+6	2.9391 E+6	3.8504 E+6	3.6257 E+6	2.9412 E+6	2.8442 E+6
27	2.9434 E+6	3.9462 E+6	3.2195 E+6	4.3251 E+6	3.9843 E+6	3.3627 E+6	2.9444 E+6
28	3.0970 E+6	6.0172 E+6	3.5350 E+6	6.0221 E+6	5.8713 E+6	3.5311 E+6	2.9506 E+6
29	3.2705 E+6	6.0823 E+6	3.6380 E+6	7.5846 E+6	6.4000 E+6	4.4641 E+6	3.3701 E+6
30	3.5309 E+6	3.8005 E+6	4.8414 E+6	1.0711 E+7	8.4900 E+6	5.8099 E+6	3.5192 E+6

10 lowest free-free elastic modes up to eigenvalue 1.44 E+6

11 lowest cantilever modes up to eigenvalue 1.44 E+6

consists of two parts, P and O. Each element has bending and extensional stiffness, and distributed mass (properties are shown in the figure). Part P consists of 48 degrees of freedom with each nodal points having X Y translational degrees of freedom and rotational degrees of freedom about the X axis. Part O consists of 32 degrees of freedom after reducing all rotational degrees of freedom by Guyan reduction [14]. To assess the accuracy of various modal representations, only part P is represented using substructure modes and then is coupled to Part O at nodal points A, B, and C through the Y and Z translational degrees of freedom.

Eigenvalues of the coupled structure, with and without reduced degrees of freedom, were used as a measure of accuracy. The calculations for eigenvalues were performed using NASTRAN Level 17.5 DMAP sequences in which the Given Eigen Solution routine was used.

Table 1 gives the eigenvalue of the elastic modes calculated with higher modes, lower modes, and both higher and lower free-free substructure modes approximated. The first column of the table shows the eigenvalues of the coupled structures with all degrees of freedom being physical. In the free-free, first order, modes approximation the 13 lowest, 13 middle, and the 13 highest elastic modes, as well as three rigid body modes, are retained with the range of substructure eigenvalues shown in the table. Equations (14), (27), and (20) are the corresponding approximations for retaining lower, middle, and higher modes, respectively. The calculated system eigenvalues close to the retained substructure eigenvalues agree well with those of the first column in Table 1. It appears that the methods developed here are capable of obtaining accurate system eigenvalues of any range.

Using the same structure, an accuracy assessment for different orders of approximation was done. The lower 10 free-free elastic modes and three rigid body modes of Part P of the structure were retained in the substructure modes representation of both the first and second-order free-free modes approximation. For the cantilever modes representation, 11 lower cantilever modes were retained. Table 2 gives the resulting system eigenvalues. The equation of coordinate transformation used for the results shown in the sixth column in the table is given by discarding the last term of equation (45). One can recognize, from Table 2, that higher order approximation methods produce more accurate results than lower order approximation methods. Although results for methods that retain lower modes, and approximate higher modes, are shown here only, better results by higher order approximations can be expected for all methods that approximate any range of substructure mode shapes.

Conclusion

Substructure mode representations considered in this paper approximate general coordinates associated with substructure modes whose eigenvalues are either significantly larger or smaller than those of the system modes of excitation. It appears that the present method approximates any range of system eigenvalues reasonably well. As suggested in Rubin's paper [11], the criterion for retaining substructure modes can be extended to one which states: "A conservative criterion for the range of component modes needed is to retain all component modes whose eigenvalues are below 2.25 (1.5^2) times the highest system eigenvalues of interest and above 40 percent of the lowest system eigenvalues."

The higher order approximations, generally, approximate system eigenvalues better than the lower order approximations with a penalty of a few more independent variables. The results of the beam structure considered in this paper show that the representation by free-free substructure modes excell those by cantilever substructure modes as long as the same range of substructure eigenvalues are retained.

Acknowledgment

The author would like to express his appreciation of Mr. Jon M. Conner, for his suggestions and comments. Also, special thanks is due to Ms. Carolyn Bateman for the typing of this paper.

References

1 Hurty, W. C., "Dynamic Analysis of Structural Systems Using Component Modes," *AIAA Journal*, Vol. 3, Apr. 1965, pp. 678-685.

2 Craig, R. R. Jr., and Bampton, M. C. C., "Coupling of Substructures for Dynamic Analysis," *AIAA Journal*, Vol. 6, July 1968, pp. 1313-1319.

3 Bamford, R., Wada, B. K., Garba, J. A., and Chisholm, J., "Dynamic

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Analysis of Large Structural Systems," *Synthesis of Vibrating Systems*, ASME booklet, Nov. 1971, pp. 57–71.

4 Coldman, R. L., "Vibration Analysis by Dynamic Partitioning," *AIAA Journal*, Vol. 7, June 1969, pp. 1152–1154.

5 Hou, S. N., "Review of Modal Synthesis Techniques and a New Approach," *Shock Vibration Bulletin*, No. 40, Part 4, 1969, pp. 25–30.

6 Dowell, E. H., "Free Vibration of an Arbitrary Structure in Terms of Component Modes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, No. 3, 1972, pp. 727–732.

7 Benfield, W. A., and Hruda, R. F., "Vibration Analysis of Structures by Component Mode Substitution," *AIAA Journal*, Vol. 9, No. 7, 1971, pp. 1255–1261.

8 Kuhar, E. J. Jr., and Stahle, C. V., "A Dynamic Transformation Method for Modal Synthesis," *AIAA Journal*, Vol. 12, No. 5, 1974, pp. 672–678.

9 Hintz, R. M., "Analytical Methods in Component Modal Synthesis," *AIAA Journal*, Vol. 13, No. 8, 1975, pp. 1007–1016.

10 MacNeal, R. H., "A Hybrid Method of Component Mode Synthesis," *Computer & Structures*, Vol. 1, Dec. 1971, pp. 581–601.

11 Rubin, S., "Improved Component-Mode Representation for Structural Dynamic Analysis," *AIAA Journal*, Vol. 13, No. 8, Aug. 1975, pp. 995–1006.

12 Craig, R. R. Jr., and Chang, C. J., "On the Use of Attachment Modes in Substructure Coupling for Dynamic Analysis," Paper 77-405, AIAA/ASME *18th Struc., Struc. Dyn., and Materials Conf.*, San Diego, Calif., 1977.

13 Klosterman, A. L., and Mclelland, W. A., "Combined Experimental and Analytical Techniques for Dynamic System Analysis," *1973 Tokyo Seminar Finite Element Analysis*, Japan.

14 Guyan, R. J., "Reduction of Stiffness and Mass Matrices," *AIAA Journal*, Vol. 3, Feb. 1965, p. 380.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Journal of Applied Mechanics

Design Data and Methods

Natural Frequencies of Out-of-Plane Vibration of Arcs

T. Irie,[1] G. Yamada,[1] and K. Tanaka[1]

The natural frequencies of out-of-plane vibration based on the Timoshenko beam theory are calculated numerically for uniform arcs of circular cross section under all combination of boundary conditions, and the results are presented in some figures.

The natural frequencies of elastic arcs were usually calculated by the classical beam theory in which the rotatory inertia and shear deformation are not considered. The numerical results obtained by the theory cannot present accurate values for arcs of considerably thick cross section, as indicated by some authors (for example reference [1]). Recently, the free out-of-plane vibration of arcs has been studied by several authors [2–7], based on the Timoshenko beam theory in which both the rotatory inertia and shear deformation are taken into account. However, there have not been sufficient data for practical use.

This Design Data paper presents the natural frequencies (the dimensionless frequency parameters) of out-of-plane vibration for uniform arcs of circular cross section with free, hinged, and clamped edge conditions. Consider a uniform arc of radius of curvature of the neutral axis R. With the angular coordinate denoted by θ and with the opening angle by α, the X, Y, and Z-axes are taken in radial, transverse, and tangential directions, respectively. On the assumption that the shear center of the cross section coincides with the centroid, the Timoshenko equations of the arc are written as

$$\frac{dQ_y^*}{Rd\theta}+\rho A\omega^2 v^*=0$$

$$-\frac{dM_x^*}{Rd\theta}+\frac{1}{R}T^*-Q_y^*+\rho I_x\omega^2\varphi=0$$

$$\frac{dT^*}{Rd\theta}+\frac{1}{R}M_x^*+\rho J_z\omega^2\psi=0 \qquad (1)$$

where ρ is the mass per unit volume, A is the cross-sectional area, I_x and J_z, respectively, are the second moment and polar moment of area of the arc, and ω is the angular frequency. The bending moment M_x^*, the torsional moment T^*, and the shearing force Q_y^*, respectively, are given by

$$M_x^*=\frac{EI_x}{R}\left(-\psi-\frac{d\varphi}{d\theta}\right),\quad T^*=\frac{GC_z}{R}\left(-\varphi+\frac{d\psi}{d\theta}\right)$$

$$Q_y^*=\kappa GA\left(\varphi+\frac{dv^*}{Rd\theta}\right) \qquad (2)$$

in terms of the transverse deflection v^*, the angle of rotation φ due to pure bending, and the angle of torsion ψ. The variables φ, v^*, ψ, M_x^*, Q_y^*, and T^* are defined to be of positive sign in the X, Y, and Z-directions. The quantity E is Young's modulus, G is the shear modulus, and C_z is the St. Venant torsional constant of the cross section. The parameter κ is the numerical factor depending on the shape of the cross section, which is 0.85 for rectangular cross section and 0.89 for circular cross section for an arc of Poisson's ratio $\nu=0.3$.

The boundary conditions of the arc are written as

$$\begin{aligned} M_x^* &= Q_y^* = T^* = 0 \quad \text{at free edge}\\ v^* &= M_x^* = T^* = 0 \quad \text{at hinged edge}\\ \varphi &= v^* = \psi = 0 \quad \text{at clamped edge} \end{aligned} \qquad (3)$$

Upon eliminating the variables M_x^*, Q_y^*, and T^*, (1) and (2) are written as

$$\kappa\frac{G}{E}s_x^2v''+\lambda^2v+\kappa\frac{G}{E}s_x^2\varphi'=0$$

$$-\kappa\frac{G}{E}s_x^2v'+(1+\mu)\psi'+\varphi''-\left(\mu+\kappa\frac{G}{E}s_x^2-\lambda^2\frac{1}{s_x^2}\right)\varphi=0$$

$$\mu\psi''-\left\{1-\lambda^2\left(\frac{1}{s_x^2}+\frac{1}{s_y^2}\right)\right\}\psi-(1+\mu)\varphi'=0 \qquad (4)$$

where, for simplicity of the analysis, the following dimensionless variables have been introduced:

$$v=\frac{1}{R}v^*,\quad (M_x,T)=\frac{R}{EI_x}(M_x^*,T^*),\quad Q_y=\frac{R^2}{EI_x}Q_y^*$$

$$s_x^2=\frac{AR^2}{I_x},\quad s_y^2=\frac{AR^2}{I_y},\quad \mu=\frac{GC_z}{EI_x}$$

$$\lambda^2=\frac{\rho AR^4\omega^2}{EI_x}\quad\left(\omega=\lambda\sqrt{\frac{EI_x}{\rho AR^4}}\right) \qquad (5)$$

The quantities s_x and s_y are the slenderness ratios, μ is the rigidity ratio of the arc, and λ denotes a frequency parameter.

With the variables taken as $\varphi=Ce^{p\theta}$, $v=qCe^{p\theta}$, and $\psi=rCe^{p\theta}$, (4) are written as the matrix equation,

[1] Professor, Associate Professor, and Instructor of Mechanical Engineering, respectively, Faculty of Engineering, Hokkaido University, Kita-13, Nishi-8, Sapporo, 060 Japan.

Manuscript received by ASME Applied Mechanics Division, October 1981; final revision, May 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Frequency equations of out-of-plane vibration of arcs

Free-free

$$\begin{vmatrix} T_{41} & T_{42} & T_{43} \\ T_{51} & T_{52} & T_{53} \\ T_{61} & T_{62} & T_{63} \end{vmatrix}_{(\alpha)} = 0$$

Free-hinged

$$\begin{vmatrix} T_{21} & T_{22} & T_{23} \\ T_{41} & T_{42} & T_{43} \\ T_{61} & T_{62} & T_{63} \end{vmatrix}_{(\alpha)} = 0$$

Free-clamped

$$\begin{vmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{vmatrix}_{(\alpha)} = 0$$

Hinged-hinged

$$\begin{vmatrix} T_{21} & T_{23} & T_{25} \\ T_{41} & T_{43} & T_{45} \\ T_{61} & T_{63} & T_{65} \end{vmatrix}_{(\alpha)} = 0$$

Hinged-clamped

$$\begin{vmatrix} T_{11} & T_{13} & T_{15} \\ T_{21} & T_{23} & T_{25} \\ T_{31} & T_{33} & T_{35} \end{vmatrix}_{(\alpha)} = 0$$

Clamped-clamped

$$\begin{vmatrix} T_{14} & T_{15} & T_{16} \\ T_{24} & T_{25} & T_{26} \\ T_{34} & T_{35} & T_{36} \end{vmatrix}_{(\alpha)} = 0$$

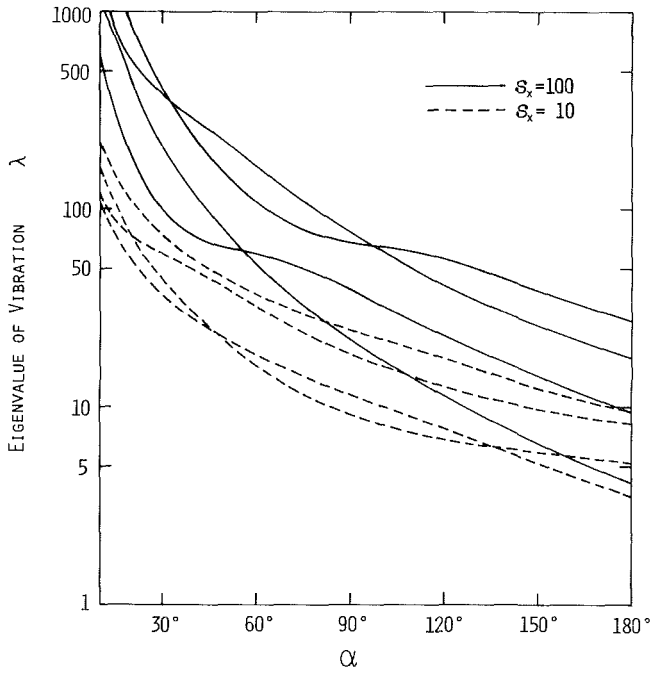


Fig. 1(*a*) Free-free arcs

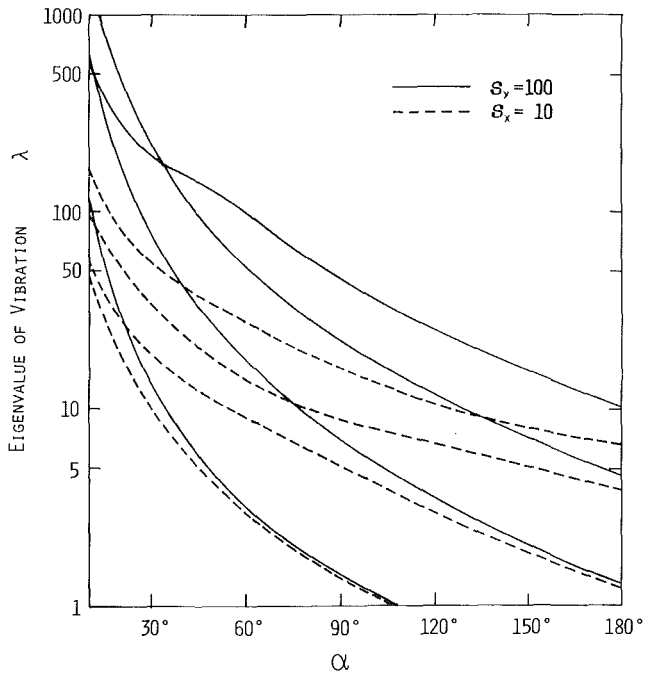


Fig. 1(*c*) Free-clamped arcs

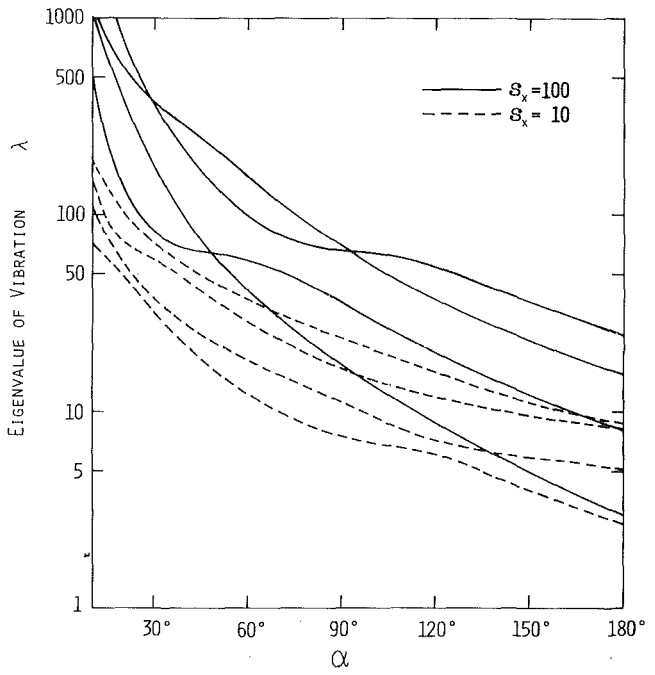


Fig. 1(*b*) Free-hinged arcs

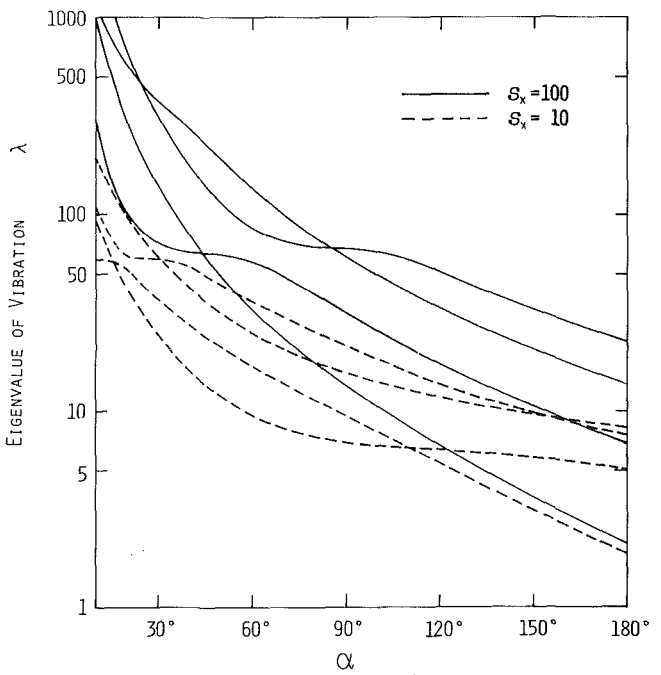


Fig. 1(*d*) Hinged-hinged arcs

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

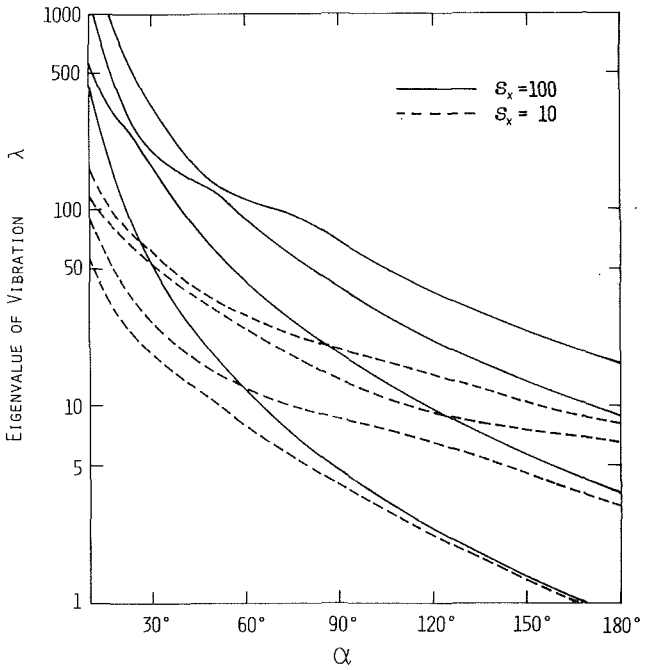


Fig. 1(e) Hinged-clamped arcs

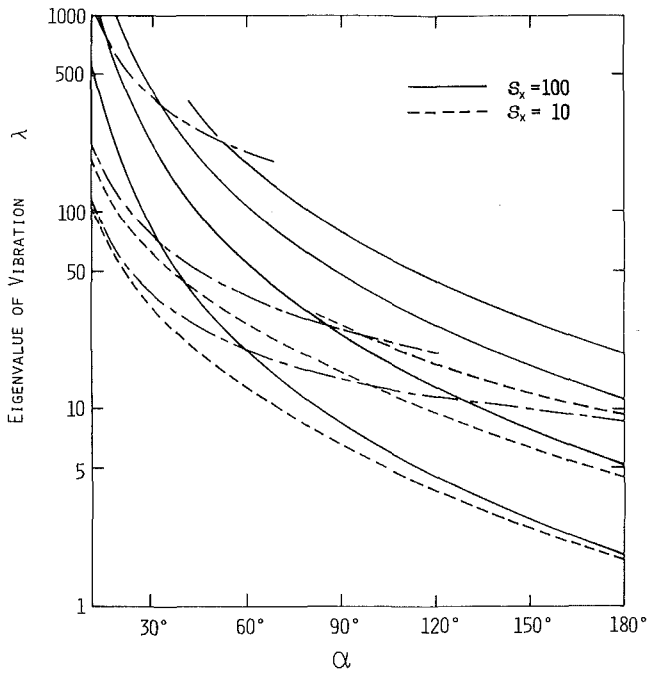


Fig. 1(f) Clamped-clamped arcs

Fig. 1 Frequency parameters λ_n of arcs of circular cross section; $\nu = 0.3$

$$\begin{bmatrix} \kappa\frac{G}{E}s_x^2 p & \kappa\frac{G}{E}s_x^2 p^2+\lambda^2 & 0 \\ p^2-\mu-\kappa\frac{G}{E}s_x^2+\lambda^2\frac{1}{s_x^2} & -\kappa\frac{G}{E}s_x^2 p & (1+\mu)p \\ -(1+\mu)p & \mu p^2-1+\lambda^2\left(\frac{1}{s_x^2}+\frac{1}{s_y^2}\right) & 0 \end{bmatrix} \begin{Bmatrix} 1 \\ q \\ r \end{Bmatrix} = 0 \qquad (6)$$

The values p are determined by calculating the eigenvalues of (6), and the values of q and r are determined by calculating the eigenvectors corresponding to the eigenvalues. The state vector $\{Z(\theta)\} = \{\varphi\ v\ \psi\ M_x\ Q_y\ T\}^T$ of the arc is expressed as

$$\{Z(\theta)\} = [T(\theta)]\{Z(\theta)\} \qquad (7)$$

by using the transfer matrix

$$[T(\theta)] = [M(\theta)][M(0)]^{-1}$$

$$[M(\theta 0)] = \begin{bmatrix} e^{p_1\theta} & \cdots & e^{p_6\theta} \\ q_1 e^{p_1\theta} & \cdots & q_6 e^{p_6\theta} \\ r_1 e^{p_1\theta} & \cdots & r_6 e^{p_6\theta} \\ -(p_1+r_1)e^{p_1\theta} & \cdots & -(p_6+r_6)e^{p_6\theta} \\ \kappa\frac{G}{E}s_x^2(1+p_1q_1)e^{p_1\theta} & \cdots & \kappa\frac{G}{E}s_x^2(1+p_6q_6)e^{p_6\theta} \\ -\mu(1-p_1r_1)e^{p_1\theta} & \cdots & -\mu(1-p_6r_6)e^{p_6\theta} \end{bmatrix} \qquad (8)$$

The substitution of (7) into a given set of the boundary conditions (3) derives the frequency equations shown in Table 1 with only the elements of $[T(\alpha)]$ necessary for the calculation.

Figures 1 present the frequency parameters of uniform arcs of circular cross section under all combination of the boundary conditions. The parameters become smaller in a wavelike manner with an increase of the opening angle α and with a decrease of the slenderness ratio $s_x = s_y$. The classical beam theory does not give accurate values for arcs of small opening angle and small slenderness ratio such as $\alpha \leq 10/s_x$ (for example, $\alpha \leq 57$ deg for $s_x = 10$) for the fundamental vibration, and for arcs of larger angle and larger slenderness ratio for higher mode vibrations, as indicated by Blevins [1] and other authors. However, the Timoshenko beam theory gives sufficiently accurate values for the frequency parameters within a wide range of the angle and the ratio of arcs. Only in clamped-clamped arcs, torsion type vibration shown by chain lines (Fig. 1(*f*)) in which the angle of torsion is dominant arises besides usual bending type vibration in which the transverse deflection is dominant [8].

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 2 The frequency parameters λ_n of clamped-clamped arcs of square and circular cross sections; $\nu = 0.3$

		Square cross section			Circular cross section		
$s_x = s_y$	n	$\alpha = 60$ deg	120 deg	180 deg	$\alpha = 60$ deg	120 deg	180 deg
20	1	16.74	4.282	1.776	16.88	4.309	1.791
	2	36.92[a]	11.69	4.982	39.70[a]	11.79	5.032
	3	40.45	22.05 [a]	10.13	40.90	22.50	10.23
	4	69.62	22.38	16.76	70.51	23.30 [a]	16.91
100	1	19.40	4.451	1.804	19.45	4.473	1.818
	2	54.03	12.83	5.198	54.14	12.89	5.242
	3	105.6	25.99	10.92	105.9	26.08	10.99
	4	172.8	43.57	18.72	173.1	43.68	18.81

[a] torsion type vibration

In Table 2, the frequency parameters of arcs of square cross section are compared with those of arcs of circular cross section under clamped-clamped edge conditions. In general, the frequency parameters of arcs of square cross section are smalller than those of arcs of circular cross section, but the difference between them is very small.

The natural frequencies of arcs are calculated numerically from (5) using the dimensionless frequency parameters presented in the Design Data.

The numerical computations presented here were carried out on a HITAC M-200H computer of the Hokkaido University Computing Center.

References

1 Blevins, R. D., *Formulas for Natural Frequency and Mode Shape*, Van Nostrand Reinhold, 1979.

2 Rao, S. S., "Effects of Transverse Shear and Rotatory Inertia on the Coupled Twist-Bending Vibrations of Circular Rings," *Journal of Sound and Vibration*, Vol. 16, No. 4, 1971, pp. 551–566.

3 Kirkhope, J., "Out-of-Plane Vibration of Thick Circular Ring," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 102, No. EM2, 1976, pp. 239–247.

4 Suzuki, K., and Takahashi, S., "Out-of-Plane Vibrations of Curved Bars Considering Shear Deformation and Rotatory Inertia," *Bulletin of the JSME*, Vol. 24, No. 193, 1981, pp. 1206–1213.

5 Davis, R., Henshell, R. D., and Warburton, G. B., "Curved Beam Finite Elements for Coupled Bending and Torsional Vibration," *Earthquake Engineering and Structural Dynamics*, Vol. 1, No. 2, 1972, pp. 165–175.

6 Irie, T., Yamada, G., and Takahashi, I., "Out-of-Plane Vibration of Arc Bar of Variable Cross-Section," *Bulletin of the JSME*, Vol. 23, No. 181, 1980, pp. 1200–1205.

7 Irie, T., Yamada, G., and Takahashi, I., "The Steady State Out-of-Plane Response of a Timoshenko Curved Beam With Internal Damping," *Journal of Sound and Vibration*, Vol. 71, No. 1, 1980, pp. 145–156.

8 Irie, T., Yamada, G., and Tanaka, K., "Free Out-of-Plane Vibration of Arcs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 439–441.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

A Simple and Effective Pipe Elbow Element—Pressure Stiffening Effects

K. J. Bathe[1] and C. A. Almeida[1]

1 Introduction

In two previous communications we presented the formulation of a simple and effective pipe elbow element, which can be employed to model the variation of ovalization along an elbow, and interaction effects between elbows of different curvatures and elbows and straight piping sections [1, 2]. The objective in this Brief Note is to show how the element formulation can also be extended in a very simple way to include internal pressure stiffening effects. As has been established experimentally and theoretically, the effect of the internal pressure on the stiffness of an elbow section can be significant when relatively thin pipes are considered. In the presentation that follows we assume that the reader is familiar with our two earlier papers, and we concentrate only on the additional evaluations necessary to include the pressure stiffening effects.

2 Additional Term in Variational Formulation

Consider the elbow in Fig. 1 subjected to an internal pressure p. As the cross section of the pressurized bend deforms due to external loading the work of the pressure acting against the change in the cross-sectional area must be considered. Thus, the potential of the pressure to be added to the variational indicator (equation (17) of reference [2]) is

$$W_{pr} = -\int_{-1}^{+1}\int_{0}^{2\pi} p\,\frac{(R - a\cos\phi)\theta}{2}\;dA(r,\phi)\,dr$$

where p is the internal pressure, $(R - a\cos\phi)$ is the longitudinal arc length of the midsurface of the bend, r is the isoparametric longitudinal coordinate, and $dA(r, \phi)$ is the differential change in the cross-sectional area of the pipe bend. This area quantity can conveniently be evaluated by using an auxiliary coordinate system $\bar{x}$, $\bar{y}$ to measure the coordinates of the points A and B on the midsurface of the bend (see Fig. 1). Let $A - B$ be the differential segment $ad\phi$ of the midsurface prior to deformation and $A' - B'$ be the same segment after deformation, then

$$dA = \frac{1}{2}\left((\bar{x}_A - \bar{x}_{B'})(\bar{y}_{A'} - \bar{y}_B) - (\bar{y}_A - \bar{y}_{B'})(\bar{x}_{A'} - \bar{x}_B)\right) \quad (1)$$

[1] Associate Professor, Mem. ASME, and Graduate Student, respectively, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

Manuscript received by ASME Applied Mechanics Division, January, 1982; final revision, April, 1982.

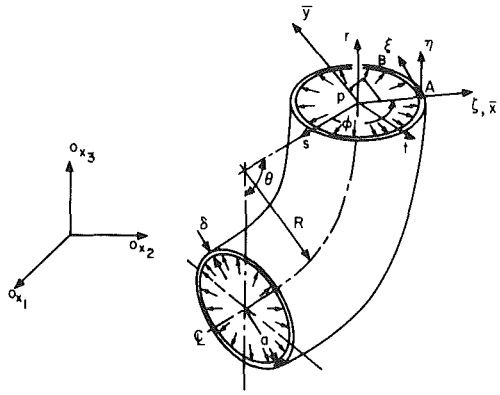


Fig. 1 Coordinate systems used in the elbow element formulation

where, with the displacements w_ξ and w_ζ into the ξ and ζ directions,

$$\begin{aligned}
&\bar{x}_A = a, \quad \bar{y}_A = 0\\
&\bar{x}_B = a\cos d\phi, \quad \bar{y}_B = a\sin d\phi\\
&\bar{x}_{A'} = a + w_\zeta, \quad \bar{y}_{A'} = w_\xi\\
&\bar{x}_{B'} = (a + w_\zeta + dw_\zeta)\cos d\phi - (w_\xi + dw_\xi)\sin d\phi\\
&\bar{y}_{B'} = (a + w_\zeta + dw_\zeta)\sin d\phi + (w_\xi + dw_\xi)\cos d\phi
\end{aligned} \quad (2)$$

Substituting from equation (2) into equation (1) we obtain

$$dA = \frac{1}{2}\Big[\underbrace{2\,a\,w_\zeta}_{\text{TERM1}} + \underbrace{w_\zeta\left(w_\zeta + \frac{dw_\xi}{d\phi}\right)}_{\text{TERM2}} + w_\xi^2 - \underbrace{w_\xi\frac{dw_\zeta}{d\phi}}_{\text{TERM3}}\Big]\,d\phi \quad (3)$$

Considering equation (3) we now note that TERM1 will not contribute to W_{pr} because the integral of w_ζ over ϕ with the ovalization displacement assumption used (see equation (8) below) is zero. TERM2 is zero because we assume that the circumferential strains vanish at the midsurface of the bend (see equation (1) of reference [1]), so that to first order

$$w_\zeta = -\frac{dw_\xi}{d\phi} \quad (4)$$

To evaluate TERM3, which is of second order, we also use that the length of the arc $A - B$ shall be equal to the length of the arc $A' - B'$ but now include second-order terms. Hence, using equation (2) we assume

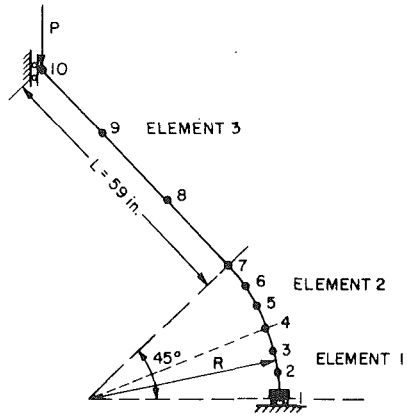
$$\begin{aligned}
(a\,d\phi)^2 = &\{dw_\zeta - (w_\xi + dw_\xi)d\phi\}^2\\
&+ \{(a + w_\zeta + dw_\zeta)d\phi + dw_\xi\}^2
\end{aligned} \quad (5)$$

and obtain with equation (4)

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

PARAMETER	BEND 1	BEND 2	BEND 3	BEND 4
R	45.	14.25	12.4	10
δ	.5	.016	.020	.016
$\bar{a}$	15.	1.5	1.75	1.5
λ	.10483	.10623	.08489	.074545

(a) Bend geometric parameters (all dimensions in inches)



Boundary Conditions:

At node 1:

—no cross-sectional rotation, $\gamma_1 = 0$.

$$-\frac{dw_\xi}{dx} = 0$$

At nodes 4 and 7:

—ovalization and derivative of ovalization continuous

(b) Finite Element Model Used

Fig. 2 Pressurized bends considered in analyses. For definition of parameters see Figs. 1 and 3.

$$w_\xi \frac{dw_\xi}{d\phi} = \frac{1}{2}\left[w_\xi^2 + \left(\frac{d^2 w_\xi}{d\phi^2}\right)^2\right] \tag{6}$$

so that

$$W_{pr} = \frac{-p}{8}\int_{-1}^{+1}\int_0^{2\pi}\left(w_\xi^2 - \left(\frac{d^2 w_\xi}{d\phi^2}\right)^2\right)(R - a\cos\phi)\,\theta\, d\phi\, dr \tag{7}$$

3 Finite Element Matrix

The finite element stiffness matrix accounting for the internal pressure effect is obtained by substituting the ovalization displacement interpolations into the expression for W_{pr}, and including the term in the total potential Π when invoking the stationarity of Π. Using the ovalization displacement interpolations of our elbow element [1, 2]

$$w_\xi(r,\phi) = \sum_{m=1}^{N_c}\sum_{k=1}^{4} h_k\, c_m^k \sin 2m\phi + \sum_{m=1}^{N_d}\sum_{k=1}^{4} h_k\, d_m^k \cos 2m\phi \tag{8}$$

where N_c and N_d are the number of ovalization displacement components to be included, we obtain (similar to the calculation of the penalty matrices in reference [2]) the following pressure stiffness matrix,

$$\mathbf{K}_{pr} = \frac{p}{2}\int_{-1}^{+1}\int_0^{2\pi}[\mathbf{G}_{p1}^T\,\mathbf{G}_{p1} - \mathbf{G}_{p2}^T\,\mathbf{G}_{p2}]\left[\frac{(R - a\cos\phi)\,\theta}{2}\right] d\phi\, dr \tag{9}$$

where

$$\mathbf{G}_{p1} = [\ldots\ldots\bar{a}_1^k\ \bar{a}_2^k\ \bar{a}_3^k\ \bar{b}_1^k\ \bar{b}_2^k\ \bar{b}_3^k\ldots\ldots] \tag{10}$$

$$\bar{a}_m^k = -(2m)^2 h_k \sin 2m\phi$$

$$\bar{b}_m^k = -(2m)^2 h_k \cos 2m\phi$$

$$\mathbf{G}_{p2} = [\ldots\ldots\tilde{a}_1^k\ \tilde{a}_2^k\ \tilde{a}_3^k\ \tilde{b}_1^k\ \tilde{b}_2^k\ \tilde{b}_3^k\ldots\ldots] \tag{11}$$

$$\tilde{a}_m^k = h_k \sin 2m\phi$$

$$\tilde{b}_m^k = h_k \cos 2m\phi$$

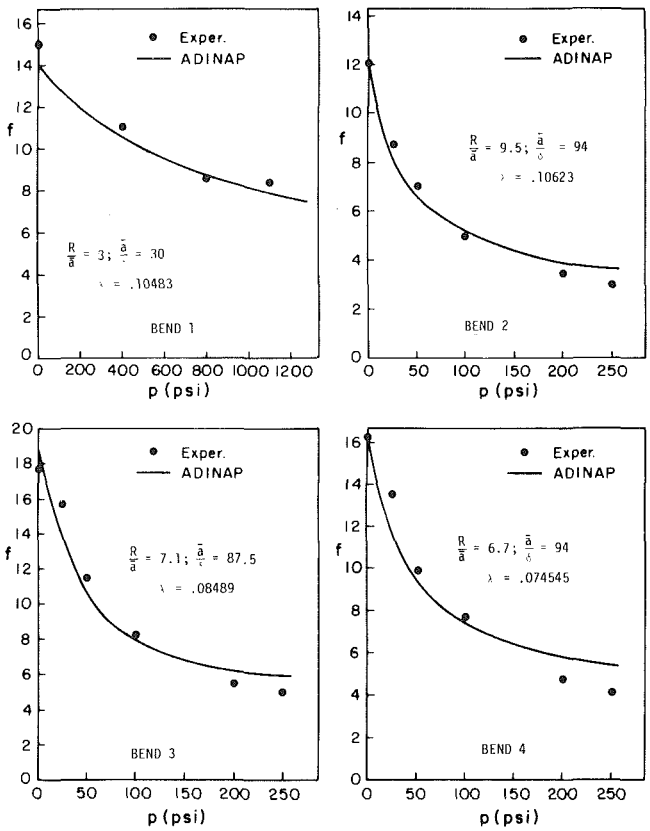


Fig. 3 Stiffening effects on the pipes of Fig. 2 due to internal pressure; $f\dfrac{4\gamma_7}{M}\dfrac{a^3 E\delta}{R}$;

γ_7 = cross-sectional rotation at node 7; $\bar{a} = a + \delta/2$; $\lambda = R\delta/(a^2\sqrt{1-\nu^2})$; $\nu = 0.3$; E and ν are Young's modulus and Poisson's ratio.

and $\mathbf{K}_{pr}$ is defined corresponding to the ovalization degrees of freedom

$$\mathbf{U}^T = [\ldots\ldots c_1^k\ c_2^k\ c_3^k\ d_1^k\ d_2^k\ d_3^k\ldots\ldots] \tag{12}$$

Referring to equations (9)–(11) we note that the evaluation of $\mathbf{K}_{pr}$ is quite inexpensive because only relatively simple integrations need be performed and the terms in $\mathbf{G}_{p1}$ and $\mathbf{G}_{p2}$ are similar.

The complete equilibrium equations corresponding to an elbow element are then

$$(\mathbf{K}_L + \mathbf{K}_p + \mathbf{K}_{pr})\,\mathbf{U} = \mathbf{R} \tag{13}$$

where $\mathbf{K}_L$ is the stiffness matrix of the elbow without internal pressure and continuity effects, as defined in reference [1], $\mathbf{K}_p$ is the penalty matrix to account for continuity effects, as defined in reference [2], and $\mathbf{R}$ is the vector of external nodal point loads [1].

4 Sample Analyses

To indicate the applicability of the foregoing analysis procedure we present the results obtained in the analyses of the bends described in Fig. 2. As shown, in these analyses three elbow elements (2 curved and 1 straight) were used, and continuity conditions were imposed between the curved and the straight sections. Figure 3 shows the calculated flexibilities for the bends and experimental results reported in references [3, 4]. It is noted that the flexibilities computed in this study are quite close to the experimental values except for higher pressures. The discrepancy at larger pressures is probably due to neglecting in the elbow formulation the midsurface strains in the ξ – direction (assumption 3, Section 2.1 of reference [1]).

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

1 Bathe, K. J., Almeida, C. A., "A Simple and Effective Pipe Elbow Element—Linear Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 93-100.

2 Bathe, K. J., Almeida, C. A., "A Simple and Effective Pipe Elbow Element—Interaction Effects," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 165-171.

3 Kafka, P. G., Dunn, M. B., "Stiffness of Curved Circular Tubes With Internal Pressure," ASME JOURNAL OF APPLIED MECHANICS, Vol. 78, 1956, pp. 247-254.

4 Rodabaugh, E. C., and George, H. H., "Effect of Internal Pressure on Flexibility and Stress-Intensification Factors of Curved Pipe or Welding Elbows," ASME JOURNAL OF APPLIED MECHANICS, Vol. 79, 1957, pp. 939-948.

Optimization of Reinforcement for a Class of Openings in Plate Structures

S. K. Dhir[1]

A numerical/analytical procedure has been developed which yields the optimum amount of reinforcement for a given hole shape in a large elastic plate under prescribed boundary stresses. This procedure is based on determining the usual two complex potentials which describe the entire stress field, constructing the strain energy density function in terms of the unknown amount of reinforcement, integrating this function around the opening boundary, and finally minimizing this integral with respect to the reinforcement. The method is first developed for a general hole shape and then demonstrated in some detail for a circular and a square-like opening.

Introduction

The use of reinforcements at the boundaries of openings is standard practice in ship and aircraft construction. The amount of reinforcement to be used is determined, in most cases, by somewhat arbitrary decisions regarding the percentage reduction it produces in the boundary stress maximums. Although intrinsically there is nothing wrong with this approach, it seems to lack a rational basis or criterion for helping the structural designer select a specific amount of reinforcement. The procedure presented in this paper is an effort to provide such a basis. Rather than determining the stress field corresponding to a given geometry of the opening and of the reinforcement, this procedure seeks to determine the reinforcement that minimizes a certain meaningful integral related to the boundary strain energy. In this way it is an inverse elasticity problem. Some investigators [1-3], in studying the related problem of optimizing unreinforced notch shapes in plates, have concluded that uniform tangential stress at the notch boundary would, in general, lead to the smallest stress concentrations. Intuitively, it appears justified to assume that in the case of reinfored notch boundaries the requirement of uniformity of boundary stresses and/or stress related quantities, such as strain energy etc., at the notch boundary would lead to more desirable designs. The optimization rationale used in the present procedure is based on this argument.

In general, stress analysis of noncircular opening reinforced with a thin member of uniform cross section is very difficult because it requires the satisfaction of a boundary condition that contains an irrational term. However, the use of MACSYMA (a symbolic manipulation language developed at MIT and in use at DTNSRDC) makes it possible to solve such problems, since a larger number of terms can be retained and manipulated in various expansions without losing track of them in the enormously long and complex algebraic expressions.

[1] David W. Taylor Naval Ship R&D Center, Bethesda, Maryland 20084.

Manuscript received by ASME Applied Mechanics Division, November, 1981; final revision, February, 1982.

The general form of the boundary condition used in this method is developed in reference [4]. The special cases of reinforcement of a circular opening and a square-like opening are discussed in greater detail and actual numerical results are included.

This paper is a logical extension of the work described in reference [5] which dealt with the optimization of the shape of a class of unreinforced openings in large plates.

Mathematical Preliminaries

An opening of general shape in a large elastic plate of isotropic material of unit thickness is reinforced by a thin member of cross-sectional area A capable of withstanding axial forces only. Then, if the opening can be mapped into a unit circle by a function $z(\zeta)$, the equivalence of complex forces between those in the reinforcement and those in the plate at the opening boundary ($\zeta = \sigma$) is given by [4]

$$\phi(\sigma) + \frac{z(\sigma)}{\overline{z'(\sigma)}}\overline{\phi'(\sigma)} + \overline{\psi(\sigma)} = \sigma P\sqrt{\frac{z'(\sigma)}{\overline{z'(\sigma)}}} + C \tag{1}$$

where $\phi(\sigma)$ and $\psi(\sigma)$ are the values of the functions $\phi(\zeta)$ and $\psi(\zeta)$ at the opening boundary, P is the axial force in the reinforcement, σ equals $e^{i\beta}$, and C is an arbitrary constant. Equivalence of tangential strain at the boundary requires

$$P = A(\sigma_\beta - \nu\sigma_\alpha) \tag{2}$$

and the elastic equilibrium of a boundary element requires [4]

$$\sigma_\alpha = P\sqrt{z'\bar{z}'}$$
$$\tau_{\alpha\beta} = -\frac{1}{\sqrt{z'\bar{z}'}}\frac{\partial P}{\partial \beta} \tag{3}$$

where the argument σ of the function z has been omitted for brevity and prime indicates differentiation with respect to σ.

The well-known relations between the stresses, σ_α, σ_β, and $\tau_{\alpha\beta}$ and the functions, $\phi(\zeta)$ and $\psi(\zeta)$, and equations (2) and (3) can be used to show that

$$P\sqrt{z'\bar{z}'}[A(1+\nu) + \sqrt{z'\bar{z}'}] = 2A(\phi'\bar{z}' + \bar{\phi}'z') \tag{4}$$

Equation (1) can be modified to

$$(\phi + \bar{\psi})\bar{z}' + z\bar{\phi}' = \sigma P\sqrt{z'\bar{z}'} \tag{5}$$

where the argument σ has been omitted, for brevity, for the ϕ, ψ, and z functions; prime indicates differentiation with respect to the argument σ; and a bar (-) represents the complex conjugate of the function. Since P is a real quantity, it is possible to represent

$$P\sqrt{z'\bar{z}'} = \sum_{n=0} c_n\left(\sigma^n + \frac{1}{\sigma^n}\right) \tag{6}$$

where the c_n are real. Functions $\phi(\zeta)$ and $\psi(\zeta)$ are known to have the following form:

$$\phi(\zeta) = S\zeta + \sum \frac{a_n}{\zeta^n}$$
$$\psi(\zeta) = D\zeta + \sum \frac{b_n}{\zeta^n} \tag{7}$$

where

$$S = \frac{p+q}{4}$$
$$D = -\frac{p-q}{2}e^{-2i\theta}$$

and p and q are uniform stresses at infinity at an angle θ to the x and y-axes, respectively. The mapping function $z(\zeta)$ can be conveniently represented by

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

1 Bathe, K. J., Almeida, C. A., "A Simple and Effective Pipe Elbow Element—Linear Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 93–100.

2 Bathe, K. J., Almeida, C. A., "A Simple and Effective Pipe Elbow Element—Interaction Effects," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 165–171.

3 Kafka, P. G., Dunn, M. B., "Stiffness of Curved Circular Tubes With Internal Pressure," ASME JOURNAL OF APPLIED MECHANICS, Vol. 78, 1956, pp. 247–254.

4 Rodabaugh, E. C., and George, H. H., "Effect of Internal Pressure on Flexibility and Stress-Intensification Factors of Curved Pipe or Welding Elbows," ASME JOURNAL OF APPLIED MECHANICS, Vol. 79, 1957, pp. 939–948.

Optimization of Reinforcement for a Class of Openings in Plate Structures

S. K. Dhir[1]

A numerical/analytical procedure has been developed which yields the optimum amount of reinforcement for a given hole shape in a large elastic plate under prescribed boundary stresses. This procedure is based on determining the usual two complex potentials which describe the entire stress field, constructing the strain energy density function in terms of the unknown amount of reinforcement, integrating this function around the opening boundary, and finally minimizing this integral with respect to the reinforcement. The method is first developed for a general hole shape and then demonstrated in some detail for a circular and a square-like opening.

Introduction

The use of reinforcements at the boundaries of openings is standard practice in ship and aircraft construction. The amount of reinforcement to be used is determined, in most cases, by somewhat arbitrary decisions regarding the percentage reduction it produces in the boundary stress maximums. Although intrinsically there is nothing wrong with this approach, it seems to lack a rational basis or criterion for helping the structural designer select a specific amount of reinforcement. The procedure presented in this paper is an effort to provide such a basis. Rather than determining the stress field corresponding to a given geometry of the opening and of the reinforcement, this procedure seeks to determine the reinforcement that minimizes a certain meaningful integral related to the boundary strain energy. In this way it is an inverse elasticity problem. Some investigators [1–3], in studying the related problem of optimizing unreinforced notch shapes in plates, have concluded that uniform tangential stress at the notch boundary would, in general, lead to the smallest stress concentrations. Intuitively, it appears justified to assume that in the case of reinfored notch boundaries the requirement of uniformity of boundary stresses and/or stress related quantities, such as strain energy etc., at the notch boundary would lead to more desirable designs. The optimization rationale used in the present procedure is based on this argument.

In general, stress analysis of noncircular opening reinforced with a thin member of uniform cross section is very difficult because it requires the satisfaction of a boundary condition that contains an irrational term. However, the use of MACSYMA (a symbolic manipulation language developed at MIT and in use at DTNSRDC) makes it possible to solve such problems, since a larger number of terms can be retained and manipulated in various expansions without losing track of them in the enormously long and complex algebraic expressions.

[1] David W. Taylor Naval Ship R&D Center, Bethesda, Maryland 20084.

Manuscript received by ASME Applied Mechanics Division, November, 1981; final revision, February, 1982.

The general form of the boundary condition used in this method is developed in reference [4]. The special cases of reinforcement of a circular opening and a square-like opening are discussed in greater detail and actual numerical results are included.

This paper is a logical extension of the work described in reference [5] which dealt with the optimization of the shape of a class of unreinforced openings in large plates.

Mathematical Preliminaries

An opening of general shape in a large elastic plate of isotropic material of unit thickness is reinforced by a thin member of cross-sectional area A capable of withstanding axial forces only. Then, if the opening can be mapped into a unit circle by a function $z(\zeta)$, the equivalence of complex forces between those in the reinforcement and those in the plate at the opening boundary ($\zeta = \sigma$) is given by [4]

$$\phi(\sigma) + \frac{z(\sigma)}{\overline{z'(\sigma)}}\overline{\phi'(\sigma)} + \overline{\psi(\sigma)} = \sigma P\sqrt{\frac{z'(\sigma)}{\overline{z'(\sigma)}}} + C \tag{1}$$

where $\phi(\sigma)$ and $\psi(\sigma)$ are the values of the functions $\phi(\zeta)$ and $\psi(\zeta)$ at the opening boundary, P is the axial force in the reinforcement, σ equals $e^{i\beta}$, and C is an arbitrary constant. Equivalence of tangential strain at the boundary requires

$$P = A(\sigma_\beta - \nu\sigma_\alpha) \tag{2}$$

and the elastic equilibrium of a boundary element requires [4]

$$\sigma_\alpha = P\sqrt{z'\bar{z}'}$$
$$\tau_{\alpha\beta} = -\frac{1}{\sqrt{z'\bar{z}'}}\frac{\partial P}{\partial \beta} \tag{3}$$

where the argument σ of the function z has been omitted for brevity and prime indicates differentiation with respect to σ.

The well-known relations between the stresses, σ_α, σ_β, and $\tau_{\alpha\beta}$ and the functions, $\phi(\zeta)$ and $\psi(\zeta)$, and equations (2) and (3) can be used to show that

$$P\sqrt{z'\bar{z}'}[A(1+\nu) + \sqrt{z'\bar{z}'}] = 2A(\phi'\bar{z}' + \bar{\phi}'z') \tag{4}$$

Equation (1) can be modified to

$$(\phi + \bar{\psi})\bar{z}' + z\bar{\phi}' = \sigma P\sqrt{z'\bar{z}'} \tag{5}$$

where the argument σ has been omitted, for brevity, for the ϕ, ψ, and z functions; prime indicates differentiation with respect to the argument σ; and a bar (-) represents the complex conjugate of the function. Since P is a real quantity, it is possible to represent

$$P\sqrt{z'\bar{z}'} = \sum_{n=0} c_n\left(\sigma^n + \frac{1}{\sigma^n}\right) \tag{6}$$

where the c_n are real. Functions $\phi(\zeta)$ and $\psi(\zeta)$ are known to have the following form:

$$\phi(\zeta) = S\zeta + \sum \frac{a_n}{\zeta^n}$$
$$\psi(\zeta) = D\zeta + \sum \frac{b_n}{\zeta^n} \tag{7}$$

where

$$S = \frac{p+q}{4}$$
$$D = -\frac{p-q}{2}e^{-2i\theta}$$

and p and q are uniform stresses at infinity at an angle θ to the x and y-axes, respectively. The mapping function $z(\zeta)$ can be conveniently represented by

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

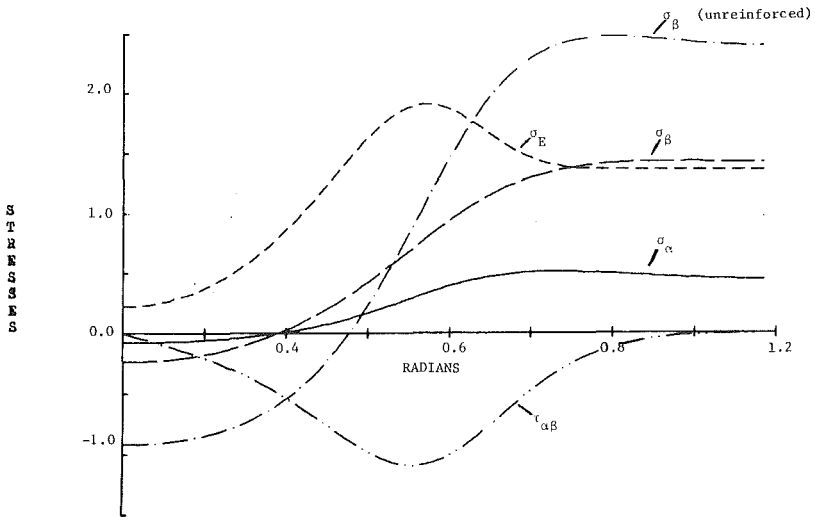


Fig. 1 Various boundary stresses around a square-like opening with optimum reinforcement

$$z(\zeta) = R\left(\zeta + \sum \frac{m_n}{\zeta^n}\right) \tag{8}$$

where m_n are known constants and R is a size factor that will be assumed unity.

The problem can now be stated in two parts: (*a*) for a given mapping function $z(\zeta)$, determine the functions $\phi(\zeta)$ and $\psi(\zeta)$ as a function of the area of reinforcement A; and (*b*) determine the value of A that minimizes a certain meaningful integral related to the boundary strain energy.

Clearly the solution of part (*a*), which forms the basis for the solution of part (*b*), would require the determination of a_n and b_n such that equations (4) and (5) are satisfied.

Substitution of equations (6), (7), and (8) into the boundary condition, equation (5), and its ensuing solution generates one set of linear simultaneous equations between a_n and c_n and another set between a_n, b_n, and c_n, both of which satisfy this boundary condition. Values of a_n in terms of c_n can be determined from the first set and substituted in equation (4) to determine c_n in terms of A by using an expansion of the irrational term $\sqrt{z'\bar{z}'}$. Taylor series expansion of this term at $\zeta = 0$ appeared to be a reasonable choice, and it was seen that, for the demonstration problem of a square-like hole, an accuracy of 1 in 10^6 could be obtained at the hole boundary by using the first eight terms after the constant term. Once the c_n are known, the two sets of simultaneous equations can, in principle, be solved to obtain a_n and b_n in terms of the unknown quantity A. The basic information for the optimization problem is then available.

Optimization

The strain energy density, V_0, of a plane elastic system is given by

$$V_0 = \frac{1}{2E}\left[\sigma_\alpha{}^2 + \sigma_\beta{}^2 - 2\nu\sigma_\alpha\sigma_\beta + 2(1+\nu)\tau_{\alpha\beta}{}^2\right] \tag{9}$$

With equations (2) and (3), this expression for V_0 can be transformed to

$$V_0 = \frac{1}{4E}\left\{\left[\frac{2}{A^2} + \frac{2(1+\nu)}{z'\bar{z}'}\right]P^2 + \frac{4(1+\nu)}{z'\bar{z}'}\left(\frac{\partial P}{\partial \beta}\right)^2\right\} \tag{9a}$$

The area of reinforcement, A, as it occurs in equation (9*a*) is a dimensionless quantity since it has been divided by the unit thickness of the plate as well as by the unit size factor R of the mapping function.

For specific loading conditions, the strain energy density, V_0, a function of β and A only, can be integrated with respect to β from 0–2π to obtain an integral, I, which can be interpreted as the strain energy in a thin region of variable thickness around the opening. This integral can now be minimized with respect to A to obtain the value of A that leads to the "most uniform" distribution of V_0 around the opening. The term "most uniform" conveys the meaning that, of all the possible parametric variations of V_0 corresponding to A, this particular distribution would have strongly attenuated peaks [5].

A distinction must be made between this integral, I, and the actual strain energy, V, in a thin slice around the opening. The strain energy itself will be obtained by integrating V_0 with respect do ds. Since ds is given by $\sqrt{z'\bar{z}'}\,d\beta$, the strain energy will be

$$V = \int V_0\sqrt{z'\bar{z}'}\,d\beta \tag{10}$$

It can be expected that a minimization of V with respect to A will lead to the "most uniform" distribution of $V_0\sqrt{z'\bar{z}'}$ rather than that of V_0 around the opening. The integral, I, was previously calculated, within a constant, in reference [5] for a square-like unreinforced opening to determine its optimum shape. For $S = 0.25$ and $D = -0.5$ the integral I was found to have a minimum value of 0.3627 at $m = -0.05$. This value of m produced a very desirable stress distribution and the lowest stress concentrations since for an unreinforced opening the integrand, V_0, of I was $\sigma_\beta{}^2$. For comparison, when this case was later repeated by minimizing V such that the integrand was $\sigma_\beta{}^2\sqrt{z'\bar{z}'}$, the minimum value of V, within the same constant was found to be 0.3705 at $m = -0.07$. The corresponding stress distribution around the opening is less desirable than when $m = -0.05$. In fact, the highest stress concentration for $m = -0.07$ was 2.505 versus 2.472 for $m = -0.05$. Nevertheless, the $\sigma_\beta{}^2\sqrt{z'\bar{z}'}$ distribution can be expected to be "most uniform" at $m = -0.07$ and the $\sigma_\beta{}^2$ distribution is the "most uniform" at $m = -0.05$. From an engineering stand point the "most uniform" $\sigma_\beta{}^2$ distribution will obviously be preferable. Other more mathematical reasons for choosing to minimize the integral of $\sigma_\beta{}^2$ are included in reference [5].

The preceding discussion justifies the use of the integral, I, rather than the strain energy, V, as the quantity to be minimized for determining the optimum A. In summary, minimization of I should, in general, be expected to lead to the "most uniform" V_0 distribution and minimization of V to the "most uniform" $V_0\sqrt{z'\bar{z}'}$ distribution. It can also be expected, as evidenced in the case of the unreinforced square

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

opening, that "most uniform" V_0 distribution would result in smaller V_0 values.

At this point it should be mentioned that consideration was also given to determining optimum A based on attenuation of other relevant quantities such as the distortion energy density (equivalent to attenuation of the octahedral stress), the dilatation energy density, the maximum shear stress, and the maximum principal stress. All these quantities except the last are either as difficult or, in fact simpler to attenuate than the strain energy density. Since the attenuation of the maximum principal stress is in general, considerably more difficult to perform (and not necessarily most desirable), the strain energy density was selected, in the case of the square-like hole, as a typical quantity to be attenuated. In the case of the circular hole other quantities were also investigated for comparison.

Examples

1. Circular Opening. In this case the irrational terms in equations (3), (4), and (5) disappear, so the manipulations are rather straightforward. Another point to be noted is that for a circular opening the strain energy, V, is the same as the Integral I. Omitting all the intermediate manipulations, the strain energy, V, can be given by

$$V = \mathrm{I} = \frac{8\pi}{E}\left[\frac{2(1+0.91A^2)S^2}{(1+1.3A)^2} + \frac{(1+11.31A^2)D^2}{(1+3.3A)^2}\right] \tag{11a}$$

the energy of distortion, V_D, or the integral of the square of octahedral stress, τ_{oct}^2, can be given by

$$V_D = \int_0^{2\pi} \frac{3.9\tau_{\mathrm{oct}}^2}{2E}\,d\beta = \frac{13.867\pi}{2E}\left[\frac{2(1-0.4A+0.79A^2)S^2}{(1+1.3A)^2} + \frac{(1-0.4A+12.79A^2)D^2}{(1+3.3A)^2}\right] \tag{11b}$$

the energy due to change in volume, V_v, can be given by

$$V_v = \frac{2.133\pi}{2E}\left[2S^2 + \frac{(1+1.3A)^2}{(1+3.3A)^2}D^2\right] \tag{11c}$$

and finally the integral, T, of the square of maximum shear stress can be given by

$$T = 16\pi\left[\frac{2(1-1.4A+0.49A^2)S^2}{(1+1.3A)^2} + \frac{(1-1.4A+16.49A^2)D^2}{(1+3.3A)^2}\right] \tag{11d}$$

The quantities V, V_D, V_v, and T were minimized to determine the optimum values of A. These computations were performed for a number of load cases as shown in Table 1. The values of A at which the minimums occurred and the corresponding minimized quantities are included in this table.

The fact that such optimum reinforcements can exist is intriguing. A popular definition of A is the ratio of area replaced as reinforcement to the area removed by the opening. Thus in the uniaxial case, if approximately 44.3 percent of the area removed by the opening is replaced by the reinforcement, the resulting stress distribution would be very desirable, i.e., one that would result in a relatively smooth boundary strain-energy density distribution. The actual stress distributions corresponding to the cases in Table 1 can be obtained using equations derived in reference [6].

2. Square Opening. The algebraic manipulations in this case are considerably more involved and lengthier than for the circular opening. A square opening described by

$$z = \sigma - \frac{0.05}{\sigma^3} \tag{12}$$

was chosen for this example because of its optimum square-like shape [5]. Solution of equations (4) and (5) and integration of equation (9) required the use of the irrational terms $(z'\,\bar{z}')^{1/2}$, $(z'\,\bar{z}')^{-1/2}$, and $(z'\,\bar{z}')^{-1}$ which were conveniently represented by Taylor series of the type

$$\sum_{n=0}^{r} G_{4n}(\sigma^{4n} + \sigma^{-4n}),$$

c.f. equation (15). In all cases truncation after eight terms could maintain an accuracy of 1 in 10^6 in the boundary values of these expansions. Equation (5) produced two sets of equations, the solution of one of which is given by

$$\left.\begin{aligned} &a_{2n-1} - \delta_{n,1}ma_1 = c_{2n} + 3ma_{2n+3} + \delta_{n,2}mS - \delta_{n,1}D \\ &a_{2n} = 0 \end{aligned}\right\};\ n \geq 1 \tag{13}$$

where $m = -0.05$. It is easy to see from equation (13) that, if either a_n or c_n is truncated, the values of individual a_n can be easily determined in terms of c_n. Because of the small value of m, the final results converged within 1 in 10^4 when the a_n beyond $n = 13$ were truncated. Once the a_n were determined, the b_n could be determined using the other set of equations. However, this procedure was not necessary since the alternative procedure for accomplishing the solution, described in the next paragraph, was more direct.

The a_n and Taylor expansion of $\sqrt{z'\bar{z}'}$ were substituted in equation (4) to obtain c_n, given by

$$\left.\begin{aligned} &c_{2n-1} = 0; \quad n \geq 1 \\ &c_{4n} = \left(\sum_{k=1}^{5} p_{4n,k}A^k\right)S \Big/ \sum_{k=0}^{5} q_k A^k; \quad 4 \geq n \geq 0 \\ &c_{4n-2} = \left(\sum_{k=1}^{4} p_{(4n-2),k}A^k\right)D \Big/ \sum_{k=0}^{4} r_k A^k; \quad 4 \geq n \geq 1 \end{aligned}\right\} \tag{14}$$

where p_n, q_n, and r_n are known coefficients but are not included here.

Once the c_n were obtained, the problem was basically solved as a function of the area of reinforcement A. The c_n were substituted in equation (6) to obtain $P\sqrt{z'\bar{z}'}$ which, in turn, was substituted in equation (9) to obtain V_0. The stresses σ_α, σ_β, and $\tau_{\alpha\beta}$ could be obtained using equations (2) and (3). In order to integrate V_0, the following Taylor expansions were used:

$$\left.\begin{aligned} (z'\bar{z}')^{-1/2} &= \sum_{n=0}^{8} L_{4n}(\sigma^{4n} + \sigma^{-4n}) \\ &\cdot \\ (z'\bar{z}')^{-1} &= \sum_{n=0}^{8} M_{4n}(\sigma^{4n} + \sigma^{-4n}) \end{aligned}\right\} \tag{15}$$

These expansions, with the now known $P\sqrt{z'\bar{z}'}$, yielded

$$\left.\begin{aligned} P &= \sum_{n=0}^{8} P_{2n}\cos 2n\beta \\ P(z'\bar{z}')^{-1/2} &= \sum_{n=0}^{8} Q_{2n}\cos 2n\beta \end{aligned}\right\} \tag{16}$$

which were used to obtain the integral

$$I = \int_0^{2\pi} V_0\,d\beta = \frac{2I_1}{A^2} + 2(1-\nu^2)I_2 + 4(1+\nu)I_3 \tag{17}$$

where

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Optimum values of reinforcement for a circular hole subject to various types of loads

Attenuated quantity →	Strain energy density		Distortion energy density		Volume energy density		Max shear stress squared	
Load cases	V_0	A	V_D	A	V_v	A	T	A
Isotropic $S=0.5; D=0$	$\frac{2.8\pi}{2E}$	*1.4286*	$\frac{1.733\pi}{2E}$	1.4286	$\frac{1.067\pi}{2E}$	−*	0	1.4286
Uniaxial $S=0.25; D=-0.5$	$\frac{3.074\pi}{2E}$	0.4429	$\frac{2.584\pi}{2E}$	0.4033	$\frac{0.349\pi}{2E}$	∞**	2.717π	0.3422
Pure shear $S=0; D=-1$	$\frac{8.15\pi}{2E}$	0.2918	$\frac{7.072\pi}{2E}$	0.2602	$\frac{0.331\pi}{2E}$	∞**	8π	0.2127

* indicates that v_v is independent of A
** infinity refers to a rigid reinforcement

Table 2 Minimized values of I at optimum values of A for a square opening

A	$\frac{2E}{16\pi}.I$
0.000001	0.36268
0.1	0.24854
0.2	0.20936
0.3	0.19546
0.4	0.19199
0.412	0.19195
0.5	0.19324
0.6	0.19671
0.7	0.20120
0.8	0.20610
0.9	0.21109
1.0	0.21600

$$\left.\begin{aligned} I_1 &= \frac{\pi}{2E}\left(P_0^2 + \frac{1}{2}\sum_{n=1}^{8} P_{2n}^2\right) \\ I_2 &= \frac{\pi}{2E}\left(Q_0^2 + \frac{1}{2}\sum_{n=1}^{8} Q_{2n}^2\right) \\ I_3 &= \frac{\pi}{2E}\left(M_0\sum_{n=1}^{8}(2nP_{2n})^2 + \sum_{n=1}^{8} 2n^2 M_{4n}P_{2n}^2 + \sum_{n=1}^{8} M_{4n}R_{4n}\right) \end{aligned}\right\} \quad (18)$$

The R_n, given as sums of products $P_m P_n$, are coefficients of $\cos n\beta$ in the expression $(\partial P/\partial\beta)^2$. The minimum value of I, which in equation (17) is a function of A, S, and D only, was determined by simply evaluating I at various values of A for a selected set of values of S and D. In fact, the three cases considered were the isotropic case ($S = 0.5$, $D = 0$), the uniaxial case ($S = 0.25$, $D = -0.5$), and the pure shear case ($S = 0$, $D = -1$). The corresponding A's at which the minimum values of I occurred were 0.90, 0.41, and 0.29. These minimum values of I could be directly compared to those of reference [5], if they were first multiplied by $2E/16\pi$. The I values for these three cases were found to be 0.2453, 0.1920, and 0.4771, respectively. It should be noted that since the value of m for all three cases was -0.05, the initially unreinforced shape is optimized only for the uniaxial case [5]. The minimization of I with respect to A for all three cases demonstrates that there is also an optimum amount of reinforcement for those openings which do not initially have an optimum shape. Table 2 gives the minimized values of I, within the constant, and the corresponding values of A for the uniaxial case only.

As expected, the value of $2E/16\pi\ I$ corresponding to $A = 0.00001 \approx 0$ is the same as that of reference [5] for $A = 0$ (unreinforced). Table 2 also shows the variation of I with respect to A and its minimum point at $A = 0.412$.

Figure 1 shows the stresses σ_α, σ_β, and $\tau_{\alpha\beta}$ for a reinforced square-like opening ($m = -0.05$) with $A = 0.412$ and also, for comparison, the equivalent stress, σ_E, based on the strain energy theory of failure and the value of σ_β for the unreinforced opening subjected to uniaxial loading. The maximum value, 1.43, of σ_β in the reinforced case occurs approximately at $\beta = 80$ deg and in the unreinforced case this maximum 2.47, occurs at approximately $\beta = 70$ deg.

Since the amount of area removed is not a constant by definition, in the case of a square opening with rounded corners, the percentage of area replaced can not be a constant either. For convenience, however, it can be referred to the half-width of the square. Thus the value of A should be divided by $1 + m$, i.e., 0.95, and multiplied by 100 to obtain the reinforcement expressed as a percentage of the area removed.

Conclusions

1. It has been shown that for openings in plates an optimum amount of reinforcement exists which corresponds to a minimum value of a strain-energy related integral.
2. It can be conjectured that the values of strain energy densities around a reinforced opening are relatively smooth when this strain-energy related integral is a minimum.
3. The actual optimum values of reinforcement for a circular opening were found to be 142, 44.3, and 29.2 percent of the area removed by the opening for isotropic, uniaxial, and pure shear loadings, respectively.
4. The optimum values of reinforcement for a square-like opening were found to be 94.7, 43.4, and 30.5 percent of the area removed by the opening for isotropic, uniaxial, and pure shear loadings, respectively.

References

1 Neuber, H., "Zur-Optimierung der Spannungskonzentration," *Continuum Mechanics and Related Problems of Analysis*, Moskau, Nauka, 1972, pp. 375-380.
2 Wheeler, L., "On the Role of Constant-Stress Surfaces in the Problem of

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Minimizing Elastic Stress Concentration," *International Journal of Solids and Structures*, Vol. 12, 1976, pp. 779–789.

3 Schnack, E., "An Optimization Procedure for Stress Concentration by the Finite Element Technique," *International Journal for Numerical Methods in Engineering*, Vol. 14, 1979, pp. 115–124.

4 Dhir, S. K., "A Hybrid Problem in Plane Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 714–716.

5 Dhir, S. K., "Optimization in a Class of Hole Shapes in Plate Structures," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, Dec. 1981, pp. 905–908.

6 Radok, J. R. M., "Problems of Plane Elasticity for Reinforced Boundaries," ASME JOURNAL OF APPLIED MECHANICS, Vol. 77, 1955, pp. 249–254.

Stability of the Vibrations of a Parameter-Excited System

M. Gürgöze[1]

The subject of this Brief Note is to introduce a special parameter-excited system with two degrees of freedom, which has its principal instability region at $\omega \approx \omega_0$ instead at $\omega \approx 2\omega_0$. It consists of a horizontal plate with a fixed pin and a slotted rigid bar on it, in which the bar can rotate and translate with respect to the plate. The stability of its vibrations is investigated for the case in which the plate is harmonicaly excited in its own plane. Starting with the Mathieu differential equation, which governs the rotational vibrations, it is possible to predict the excitation frequencies, which must be avoided, to ensure the stability.

Nomenclature

0	=	fixed pin on the plate
C	=	center of mass of the bar
ϕ	=	the rotational angle of the bar with respect to the plate
u	=	the distance between the points 0 and C
a	=	the distance of the springs and dashpots from point 0
$c/2$	=	the damping coefficient of each dashpot, which is assumed to be very small
$k/2, \bar{k}/2$	=	spring constants of the springs, where $\bar{k}/2$ is assumed to be small, so that $\bar{\omega}^2 = \bar{k}/m$ is small compared with the square of the excitation frequency (i.e., $\omega^2 > \bar{\omega}^2$)
m	=	mass of the bar
J_c, i_c	=	the moment of inertia of the bar with respect to the perpendicular axis to the bar at point C and the corresponding radius of gyration

Introduction

The system shown in Fig. 1 consists of a horizontal plate and a slotted rigid bar placed on it, so that the bar can rotate and translate with respect to the plate. While two springs and dashpots, which have one end fixed on the plate, act against the rotational motion, two soft springs act against the translational motion of the bar.

Formulation of the Problem

Let us assume that the plate is set in a simple harmonic motion in its own horizontal plane according to:

$$s = S\cos\omega t, \tag{1}$$

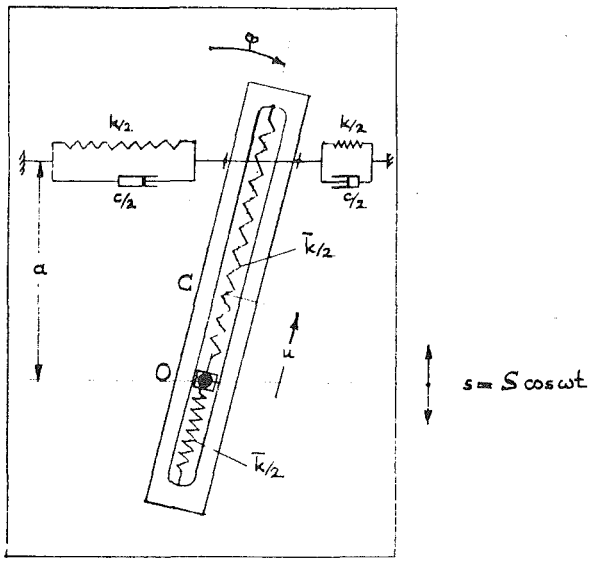


Fig. 1 Parameter-excited system with two degrees of freedom

where ω is the excitation frequency and S is the amplitude of the motion.

Neglecting the frictional effects, assuming that the rotational vibrations are small,

$$\sin\phi \approx \phi, \quad \cos\phi \approx 1 \tag{2}$$

and furthermore that 0 and C coincide in the equilibrium position, then the following equations of motion are obtained:

$$(J_c + mu^2)\ddot{\phi} + 2mu\dot{u}\dot{\phi} + ca^2\dot{\phi} + mS\omega^2 u\phi\cos\omega t + ka^2\phi = 0 \tag{3}$$

$$\ddot{u} - \dot{\phi}^2 u + \bar{\omega}^2 u - \omega^2 S\cos\omega t = 0, \tag{4}$$

where the dot represents the differentiation with respect to time and $\bar{\omega}$ is the eigenfrequency of the translational motion of the bar. One obtains the particular solution of equation (4) as:

$$\phi = 0, \quad u = U_p\cos\omega t \tag{5}$$

where

$$U_p = \omega^2 S/(\bar{\omega}^2 - \omega^2). \tag{6}$$

After linearization about this particular solution, essentially the following rheolinear differential equation is obtained:

$$\ddot{\phi} + [ca^2/(J_c + mU_p^2\cos^2\omega t)]\dot{\phi} + f_1(t)\dot{\phi} + f_2(t)\phi = 0 \tag{7}$$

where

$$f_1(t) = -mU_p^2\,\omega\sin 2\omega t/(J_c + mU_p^2\cos^2\omega t) \tag{8}$$

and

$$f_2(t) = (ka^2 + mSU_p\omega^2\cos^2\omega t)/(J_c + mU_p^2\cos^2\omega t). \tag{9}$$

The third term in equation (7) can be eliminated using the well-known linear transformation [1]

$$\phi(t) = \alpha(t)\ .\ \exp(-1/2\ .\ \int_0^t f_1(\tau)d\tau), \tag{10}$$

so that the equation

$$\ddot{\alpha} + [ca^2/(J_c + mU_p^2\cos^2\omega t)]\dot{\alpha} + f(t)\alpha = 0 \tag{11}$$

in the new variable α is obtained, where

$$f(t) = \frac{m\omega U_p^2}{4(J_c + mU_p^2\cos^2\omega t)^2}[(J_c + mU_p^2\cos^2\omega t)(4\omega\cos 2\omega t + \frac{4Ka^2}{m\omega U_p^2} + 4\omega\frac{S}{U_p}\cos^2\omega t) + m\omega U_p^2\sin^2 2\omega t + 2ca^2\sin 2\omega t] \tag{12}$$

[1]Assistant Professor, Faculty of Mechanical Engineering, Technical University of Istanbul, Macka-Istanbul, Turkey. Presently at Universitaet-Gesamthochschule Wuppertal, Fachbereich Sicherheitstechnik, D-5600 Wuppertal 1, Germany.

Manuscript received by ASME Applied Mechanics Division, July, 1981; final revision, May, 1982.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Minimizing Elastic Stress Concentration," *International Journal of Solids and Structures*, Vol. 12, 1976, pp. 779–789.

3 Schnack, E., "An Optimization Procedure for Stress Concentration by the Finite Element Technique," *International Journal for Numerical Methods in Engineering*, Vol. 14, 1979, pp. 115–124.

4 Dhir, S. K., "A Hybrid Problem in Plane Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 714–716.

5 Dhir, S. K., "Optimization in a Class of Hole Shapes in Plate Structures," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, Dec. 1981, pp. 905–908.

6 Radok, J. R. M., "Problems of Plane Elasticity for Reinforced Boundaries," ASME JOURNAL OF APPLIED MECHANICS, Vol. 77, 1955, pp. 249–254.

Stability of the Vibrations of a Parameter-Excited System

M. Gürgöze[1]

The subject of this Brief Note is to introduce a special parameter-excited system with two degrees of freedom, which has its principal instability region at $\omega \approx \omega_0$ instead at $\omega \approx 2\omega_0$. It consists of a horizontal plate with a fixed pin and a slotted rigid bar on it, in which the bar can rotate and translate with respect to the plate. The stability of its vibrations is investigated for the case in which the plate is harmonicaly excited in its own plane. Starting with the Mathieu differential equation, which governs the rotational vibrations, it is possible to predict the excitation frequencies, which must be avoided, to ensure the stability.

Nomenclature

0	=	fixed pin on the plate
C	=	center of mass of the bar
ϕ	=	the rotational angle of the bar with respect to the plate
u	=	the distance between the points 0 and C
a	=	the distance of the springs and dashpots from point 0
$c/2$	=	the damping coefficient of each dashpot, which is assumed to be very small
$k/2, \bar{k}/2$	=	spring constants of the springs, where $\bar{k}/2$ is assumed to be small, so that $\bar{\omega}^2 = \bar{k}/m$ is small compared with the square of the excitation frequency (i.e., $\omega^2 > \bar{\omega}^2$)
m	=	mass of the bar
J_c, i_c	=	the moment of inertia of the bar with respect to the perpendicular axis to the bar at point C and the corresponding radius of gyration

Introduction

The system shown in Fig. 1 consists of a horizontal plate and a slotted rigid bar placed on it, so that the bar can rotate and translate with respect to the plate. While two springs and dashpots, which have one end fixed on the plate, act against the rotational motion, two soft springs act against the translational motion of the bar.

Formulation of the Problem

Let us assume that the plate is set in a simple harmonic motion in its own horizontal plane according to:

$$s = S\cos\omega t, \tag{1}$$

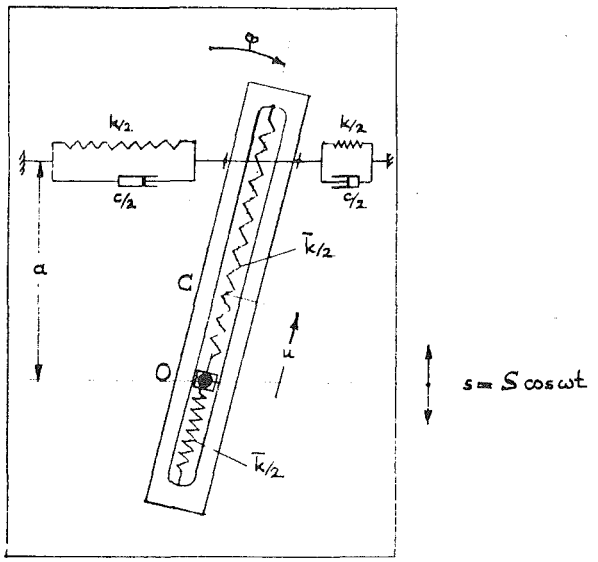


Fig. 1 Parameter-excited system with two degrees of freedom

where ω is the excitation frequency and S is the amplitude of the motion.

Neglecting the frictional effects, assuming that the rotational vibrations are small,

$$\sin\phi \approx \phi, \quad \cos\phi \approx 1 \tag{2}$$

and furthermore that 0 and C coincide in the equilibrium position, then the following equations of motion are obtained:

$$(J_c + mu^2)\ddot{\phi} + 2mu\dot{u}\dot{\phi} + ca^2\dot{\phi} + mS\omega^2 u\phi\cos\omega t + ka^2\phi = 0 \tag{3}$$

$$\ddot{u} - \dot{\phi}^2 u + \bar{\omega}^2 u - \omega^2 S\cos\omega t = 0, \tag{4}$$

where the dot represents the differentiation with respect to time and $\bar{\omega}$ is the eigenfrequency of the translational motion of the bar. One obtains the particular solution of equation (4) as:

$$\phi = 0, \quad u = U_p\cos\omega t \tag{5}$$

where

$$U_p = \omega^2 S/(\bar{\omega}^2 - \omega^2). \tag{6}$$

After linearization about this particular solution, essentially the following rheolinear differential equation is obtained:

$$\ddot{\phi} + [ca^2/(J_c + mU_p^2\cos^2\omega t)]\dot{\phi} + f_1(t)\dot{\phi} + f_2(t)\phi = 0 \tag{7}$$

where

$$f_1(t) = -mU_p^2\,\omega\sin 2\omega t/(J_c + mU_p^2\cos^2\omega t) \tag{8}$$

and

$$f_2(t) = (ka^2 + mSU_p\omega^2\cos^2\omega t)/(J_c + mU_p^2\cos^2\omega t). \tag{9}$$

The third term in equation (7) can be eliminated using the well-known linear transformation [1]

$$\phi(t) = \alpha(t)\ .\ \exp(-1/2\ .\ \int_0^t f_1(\tau)d\tau), \tag{10}$$

so that the equation

$$\ddot{\alpha} + [ca^2/(J_c + mU_p^2\cos^2\omega t)]\dot{\alpha} + f(t)\alpha = 0 \tag{11}$$

in the new variable α is obtained, where

$$f(t) = \frac{m\omega U_p^2}{4(J_c + mU_p^2\cos^2\omega t)^2}[(J_c + mU_p^2\cos^2\omega t)(4\omega\cos 2\omega t + \frac{4Ka^2}{m\omega U_p^2} + 4\omega\frac{S}{U_p}\cos^2\omega t) + m\omega U_p^2\sin^2 2\omega t + 2ca^2\sin 2\omega t] \tag{12}$$

[1]Assistant Professor, Faculty of Mechanical Engineering, Technical University of Istanbul, Macka-Istanbul, Turkey. Presently at Universitaet-Gesamthochschule Wuppertal, Fachbereich Sicherheitstechnik, D-5600 Wuppertal 1, Germany.

Manuscript received by ASME Applied Mechanics Division, July, 1981; final revision, May, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

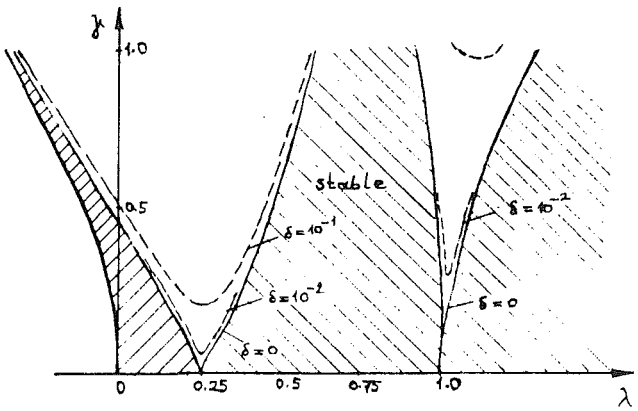


Fig. 2 Strutt-stability diagram

Transformation (10) does not change the stability behavior.

Now a nondimensional quantity κ is defined and is assumed to be much smaller than 1:

$$\kappa = mU_p^2/J_c << 1. \tag{13}$$

This assumption means physically, that the amplitude of the reciprocating motion of the plate is much smaller than the radius of gyration of the bar. Using this assumption and basic trigonometric identities, equation (11) can be simplified to the following form:

$$\ddot{\alpha} + \frac{ca^2}{J_c}\dot{\alpha} + [ka^2/J_c - \frac{\kappa}{2}(\omega^2 - \bar{\omega}^2) + \frac{\kappa}{2}(\omega^2 + \bar{\omega}^2)\cos 2\omega t]\alpha = 0, \tag{14}$$

where the damping term in the bracket is neglected in comparison with the other terms ($c\kappa << 1$).

If now a nondimensional time is defined as:

$$\tau = 2\omega t \tag{15}$$

the preceding equation takes the form of the well known "damped Mathieu differential equation" in its "normal form" [2]:

$$\alpha'' + 2\delta\alpha' + (\lambda + \gamma\cos\tau)\alpha = 0 \tag{16}$$

where the prime represents the differentiation with respect to τ and the following abbreviations are used:

$$\delta = ca^2/4\omega J_c \tag{17}$$

$$\omega_0^2 = ka^2/J_c \tag{18}$$

$$\lambda = \omega_0^2/4\omega^2 - \kappa(\omega^2 - \bar{\omega}^2)/8\omega^2 \tag{19}$$

$$\gamma = \kappa(\omega^2 + \bar{\omega}^2)/8\omega^2. \tag{20}$$

ω_0 is the eigenfrequency of the bar in the undamped case, when it is restricted to rotate only about a perpendicular axis through C.

To analyze the stability of the solutions of equation (16) one can use the Strutt-stability diagram, which is shown in Fig. 2. Considering equation (20) and the assumption stated in (13), it is seen that the values of the parameter γ lie in a region very close to the positive γ-axis of this diagram. On the other hand, it is known that the regions of instability on this diagram start at the discrete points on the λ-axis given by [2]:

$$\lambda = (n/2)^2, \quad (n = 1,2,3,\ldots). \tag{21}$$

Results and Comments

Looking at this result, one can state that the excitation frequency ω of the plate should not be close to the frequencies stated in the following, so that instability does not occur:

$$\omega = \omega_0/n, \quad (n = 1,2,3,\ldots). \tag{22}$$

Also, of course, $\omega = \bar{\omega}$ should be avoided to prevent conventional translational resonance of the bar.

It is seen that the principal instability region is at $\omega \approx \omega_0$ and not at $\omega \approx 2\omega_0$, which is usual, due to the fact that every time when 0 and C coincide, the moment of inertia forces of the bar about 0 changes its sense, so that a synchronization of the parameter excitation and the rotational motion is achieved.

It is also worth noting, that as long as the excitation frequency is greater than the rotational eigenfrequency, the small vibrations of the bar are stable. Furthermore, by the method of stability in the first approximation, the stability behavior of the particular solution (5) of the original nonlinear problem is also proved.

Acknowledgment

The author is indebted to Prof. Dr. P. C. Müller (University of Wuppertal) for his valuable contributions.

References

1 Magnus, K., *Schwingungen*, Teubner Verlag, Stuttgart, 1961.

2 Müller, P. C., and Schiehlen, W. O., *Lineare Schwingungen*, Akad. Verlagsges., Wiesbaden, 1976.

Axisymmetric Vortex Breakdown in Rotating Fluid Within a Container

H. J. Lugt[1] and H. J. Haussling[1]

Vortex breakdown was first discovered by Peckham and Atkinson [1] in 1957 in the tip vortices of delta wings and independently by the senior author [2] in 1959 in pipes. Since then a flood of papers has appeared on various theories and experimental observations, which have been reviewed in a series of publications [3–6].

Among the various forms and appearances of vortex breakdown in pipes, tornadoes, combustion chambers, and on wings, there is one case that has attracted almost no attention in the literature, and this is vortex breakdown in the rotating fluid of a closed container. It is even argued whether this phenomenon is vortex breakdown (in fact, one of the reviewers questioned this). A close look at the flow field, however, reveals that this phenomenon is indeed vortex breakdown.

Consider a circular-cylindrical container, completely filled with an incompressible fluid. The cover can rotate with angular velocity Ω coaxially to the fixed container. Before time $t = 0$ (time is made dimensionless with Ω), the cover and the fluid are at rest. At $t = 0$ the cover abruptly starts rotating with constant Ω, and a meridional and azimuthal circulation develops.

For the mathematical formulation of the problem, cylindrical polar coordinates r, ϕ, and z are introduced, where r and z are made dimensionless by the radius R of the container. The corresponding velocity components are u, v, and w made dimensionless with ΔR, in a reference frame fixed to the container. Under the restriction of axisymmetry, that is, $\partial/\partial\phi = 0$, the stream function-vorticity form of the Navier-Stokes equations for the meridional flow can be used. The dimensionless stream function ψ and the azimuthal component of the dimensionless vorticity vector ζ are defined by

$$u = \frac{1}{r}\psi_z, \quad w = -\frac{1}{r}\psi_r,$$

$$\zeta = u_z - w_r.$$

[1] David W. Taylor Naval Ship R&D Center, Bethesda, Md. 20084.

Manuscript received by ASME Applied Mechanics Division, September, 1981; final revision, April, 1982.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

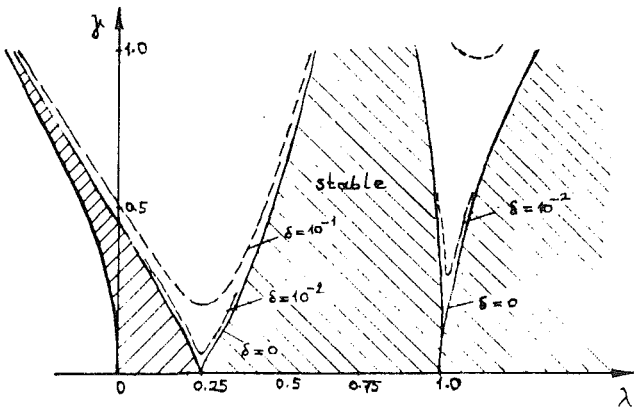


Fig. 2 Strutt-stability diagram

Transformation (10) does not change the stability behavior.

Now a nondimensional quantity κ is defined and is assumed to be much smaller than 1:

$$\kappa = mU_p^2/J_c << 1. \tag{13}$$

This assumption means physically, that the amplitude of the reciprocating motion of the plate is much smaller than the radius of gyration of the bar. Using this assumption and basic trigonometric identities, equation (11) can be simplified to the following form:

$$\ddot{\alpha} + \frac{ca^2}{J_c}\dot{\alpha} + [ka^2/J_c - \frac{\kappa}{2}(\omega^2 - \bar{\omega}^2) + \frac{\kappa}{2}(\omega^2 + \bar{\omega}^2)\cos 2\omega t]\alpha = 0, \tag{14}$$

where the damping term in the bracket is neglected in comparison with the other terms ($c\kappa << 1$).

If now a nondimensional time is defined as:

$$\tau = 2\omega t \tag{15}$$

the preceding equation takes the form of the well known "damped Mathieu differential equation" in its "normal form" [2]:

$$\alpha'' + 2\delta\alpha' + (\lambda + \gamma\cos\tau)\alpha = 0 \tag{16}$$

where the prime represents the differentiation with respect to τ and the following abbreviations are used:

$$\delta = ca^2/4\omega J_c \tag{17}$$

$$\omega_0^2 = ka^2/J_c \tag{18}$$

$$\lambda = \omega_0^2/4\omega^2 - \kappa(\omega^2 - \bar{\omega}^2)/8\omega^2 \tag{19}$$

$$\gamma = \kappa(\omega^2 + \bar{\omega}^2)/8\omega^2. \tag{20}$$

ω_0 is the eigenfrequency of the bar in the undamped case, when it is restricted to rotate only about a perpendicular axis through C.

To analyze the stability of the solutions of equation (16) one can use the Strutt-stability diagram, which is shown in Fig. 2. Considering equation (20) and the assumption stated in (13), it is seen that the values of the parameter γ lie in a region very close to the positive γ-axis of this diagram. On the other hand, it is known that the regions of instability on this diagram start at the discrete points on the λ-axis given by [2]:

$$\lambda = (n/2)^2, \quad (n = 1,2,3,\ldots). \tag{21}$$

Results and Comments

Looking at this result, one can state that the excitation frequency ω of the plate should not be close to the frequencies stated in the following, so that instability does not occur:

$$\omega = \omega_0/n, \quad (n = 1,2,3,\ldots). \tag{22}$$

Also, of course, $\omega = \bar{\omega}$ should be avoided to prevent conventional translational resonance of the bar.

It is seen that the principal instability region is at $\omega \approx \omega_0$ and not at $\omega \approx 2\omega_0$, which is usual, due to the fact that every time when 0 and C coincide, the moment of inertia forces of the bar about 0 changes its sense, so that a synchronization of the parameter excitation and the rotational motion is achieved.

It is also worth noting, that as long as the excitation frequency is greater than the rotational eigenfrequency, the small vibrations of the bar are stable. Furthermore, by the method of stability in the first approximation, the stability behavior of the particular solution (5) of the original nonlinear problem is also proved.

Acknowledgment

The author is indebted to Prof. Dr. P. C. Müller (University of Wuppertal) for his valuable contributions.

References

1 Magnus, K., *Schwingungen*, Teubner Verlag, Stuttgart, 1961.

2 Müller, P. C., and Schiehlen, W. O., *Lineare Schwingungen*, Akad. Verlagsges., Wiesbaden, 1976.

Axisymmetric Vortex Breakdown in Rotating Fluid Within a Container

H. J. Lugt[1] and H. J. Haussling[1]

Vortex breakdown was first discovered by Peckham and Atkinson [1] in 1957 in the tip vortices of delta wings and independently by the senior author [2] in 1959 in pipes. Since then a flood of papers has appeared on various theories and experimental observations, which have been reviewed in a series of publications [3–6].

Among the various forms and appearances of vortex breakdown in pipes, tornadoes, combustion chambers, and on wings, there is one case that has attracted almost no attention in the literature, and this is vortex breakdown in the rotating fluid of a closed container. It is even argued whether this phenomenon is vortex breakdown (in fact, one of the reviewers questioned this). A close look at the flow field, however, reveals that this phenomenon is indeed vortex breakdown.

Consider a circular-cylindrical container, completely filled with an incompressible fluid. The cover can rotate with angular velocity Ω coaxially to the fixed container. Before time $t = 0$ (time is made dimensionless with Ω), the cover and the fluid are at rest. At $t = 0$ the cover abruptly starts rotating with constant Ω, and a meridional and azimuthal circulation develops.

For the mathematical formulation of the problem, cylindrical polar coordinates r, ϕ, and z are introduced, where r and z are made dimensionless by the radius R of the container. The corresponding velocity components are u, v, and w made dimensionless with ΔR, in a reference frame fixed to the container. Under the restriction of axisymmetry, that is, $\partial/\partial\phi = 0$, the stream function-vorticity form of the Navier-Stokes equations for the meridional flow can be used. The dimensionless stream function ψ and the azimuthal component of the dimensionless vorticity vector ζ are defined by

$$u = \frac{1}{r}\psi_z, \quad w = -\frac{1}{r}\psi_r,$$

$$\zeta = u_z - w_r.$$

[1] David W. Taylor Naval Ship R&D Center, Bethesda, Md. 20084.

Manuscript received by ASME Applied Mechanics Division, September, 1981; final revision, April, 1982.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

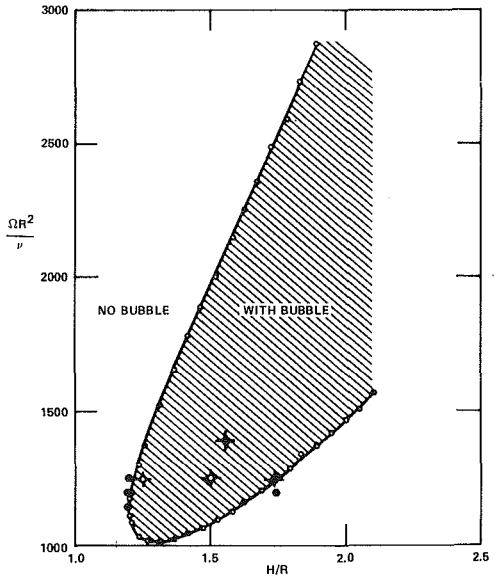


Fig. 1 Occurrence of vortex breakdown (separation bubbles). o-o-o experimentally determined dividing line (Vogel [13]); computed: ⊙ bubble present, • no bubble.

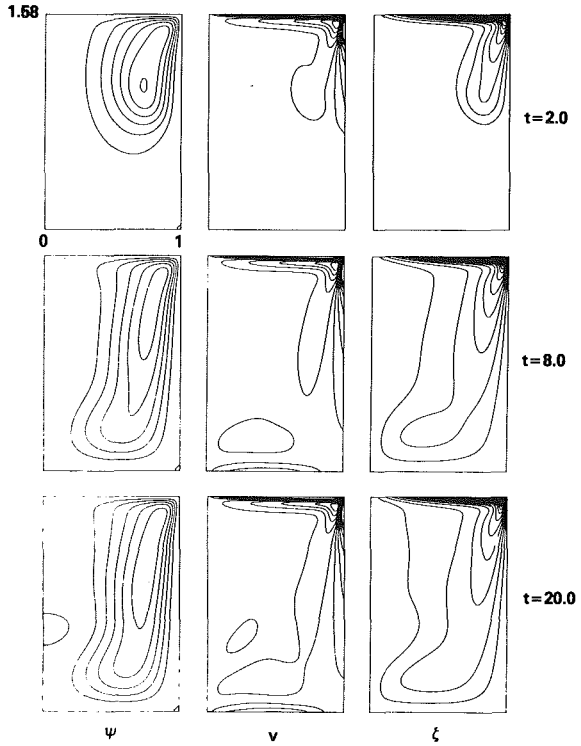


Fig. 2 Patterns of constant ψ, v and ζ for Re = 1350, δ = 1.58 at various times in the transient phase after an abrupt start of the cover

The Reynolds number is introduced as $\mathrm{Re} = \Omega R^2/\nu$, where ν is the kinematic viscosity. The geometric parameter is $\delta = H/R$ with H the height of the container.

Numerical solutions to this problem were published by the authors in [7, 8] for $\delta = 1$ and various Re. Only transient separation bubbles at the axis of rotation were observed. They vanished as $t \to \infty$ and were interpreted as transient inertial waves. In contrast, the experimental results by Vogel [9] and Ronnenberg [10] revealed steady-state separation bubbles for larger values of δ. The experimentally determined region of the (Re, δ) parameter space in which steady-state separation bubbles occur, is indicated in Fig. 1 by the shaded area.

The purpose of the new, additional numerical computations is (*a*) to analyze the time development of the flow field,

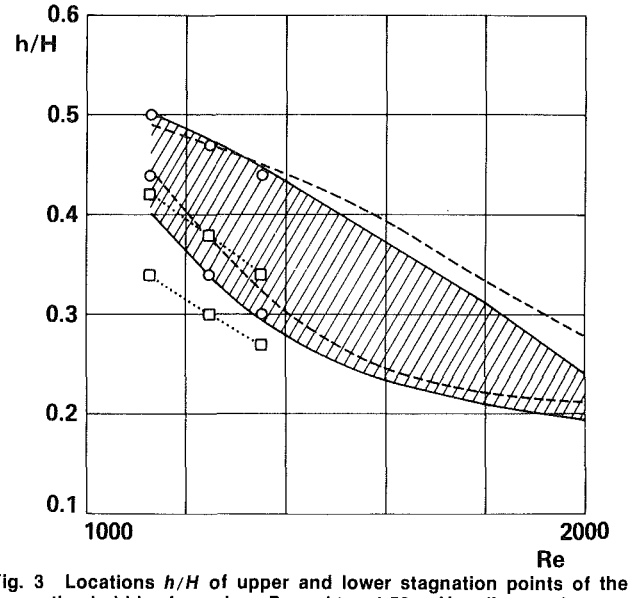


Fig. 3 Locations h/H of upper and lower stagnation points of the separation bubbles for various Re and δ = 1.58. ··· Vogel's experiments, – Ronnenberg's experiments, □ computed with 41 × 41 grid, ○ computed with 51 × 51 grid

demonstrating the character of the phenomenon as a "frozen" inertial wave, and to interpret this special kind of vortex breakdown in a simple way with reference to earlier work in the literature; and (*b*) to verify numerically the experimental data of Vogel and Ronnenberg.

(*a*) For the special case Re = 1355, δ = 1.58 chosen from Vogel's data, the flow development from the onset of motion to the final steady state is displayed in the form of lines of constant ψ, v, and ζ in Fig. 2. The abrupt start of the cover at $t = 0$ creates a flow singularity at the corner $r = 1$, $z = \delta$ from which the fluid motion starts. A meridional circulation, established through the Ekman and Stewartson layers, carries angular momentum from the outside wall region of the container toward the axis. This meridional flow exhibits monotonic or wavy streamline patterns, depending on whether the parabolic or the hyperbolic behavior of the solutions prevails [8]. In the latter case, which is displayed in Fig. 2, contraction and swelling of the upward stream around the axis occurs. The diverging streamlines and the decrease of v may retard the w-component of the velocity along the axis so much that a stagnation point appears with a backflow behind it. This abrupt enlargement of the vortex core with a separation region is typical for vortex breakdown. The divergence of the streamlines and slowing of v leads to a disturbance of the equilibrium between centrifugal acceleration and pressure gradient. A restoring force, characteristic of the hyperbolic behavior of rotating fluids, causes the fluid to converge and to close the separation region. This interpretation follows the idea of boundary-layer separation as an explanation of vortex breakdown as proposed by Gartshore [4]. The boundary-layer analogy, however, cannot explain the closing of the separation region. A further advance in the explanation was made by Grabowski and Berger [11] who described the closing as an "over-shoot" phenomenon. This, of course, points to the hyperbolic nature of the flow. In the transient period of the computed case Re = 1355, δ = 1.58 the bubble appears at $t = 10.0$ and approaches a steady state with small oscillations in the bubble shape.

The phenomenon described in the foregoing clearly has the typical characteristics of vortex breakdown. For instance, Faler and Leibovich [12] identify vortex breakdown "by the development of a stagnation point on the axis, followed by a region of reversed axial flow encapsulated by a greatly swollen stream surface."

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(b) In the steady state the locations and sizes of the separation bubbles for Re = 1130, 1250, 1355, $\delta = 1.58$ are compared with Vogel's and Ronnenberg's observations in Fig. 3. In the numerical calculations the lower separation point shifts slightly upward when the mesh is refined. Ronnenberg's laser measurements are more accurate than those in Vogel's experiments. Hence, Fig. 3 shows that the agreement between computations and experiments is satisfactory.

Good agreement is obtained for the location of the limiting curve as shown in Fig. 1. Two regions of Vogel's curve were verified as follows: the computed flow fields for points in Fig. 1 outside the region of bubble occurrence exhibit no stagnation point; inside the region tiny bubbles appear, whose resolution is a function of the grid.

References

1 Peckham, D. H., and Atkinson, S. A., *British ARC CP*, No. 508, 1957.

2 Lugt, H. J., Ph.D. Thesis, University of Stuttgart, Germany, 1959 (short version translated by "The British Hydromechanics Research Ass." Report T 716, Feb. 1962).

3 Sarpkaya, T., "Vortex Breakdown in Swirling Conicad flows," *AIAA Journal*, Vol. 9, 1971, pp. 1792-1799.

4 Hall, M. G., *Annual Review of Fluid Mechanics* (Annual Review Inc., Palo Alto, Calif.) Vol. 4, 1972, pp. 195—218.

5 Leibovich, S., *Annual Review of Fluid Mechanics* (Annual Review Inc., Palo Alto, Calif.) Vol. 10, 1978, pp. 221—246.

6 Escudier, M. P., and Zehnder, N., "Vortex-Flow regimes," *J. of Fluid Mech.*, Vol. 115, 1982, pp. 105—121.

7 Lugt, H. J., and Haussling, H. J., *IUTAM Symposium on Unsteady Boundary Layers*, Les Presses de l'Universite Laval, Quebec, 1972, pp. 1366—1396.

8 Lugt, H. J., and Haussling, H. J., *Acta Mechanica*, Vol. 18, 1973, pp. 255—272.

9 Vogel, H. U., *Max-Planck-Institut für Strömungsforschung*, Gottingen, Bericht 6, 1968, pp. 1-91.

10 Ronnenberg, B., *Max-Planck-Institut für Stömungsforchung*, Gottingen, Bericht 20, 1977, pp. 1-73.

11 Grabowski, W. J., and Berger, S. A., *J. of Fluid Mech.*, Vol. 75, 1976, pp. 525—544.

12 Faler, J. H., and Leibovich, S., *Phys. of Fluids*, Vol. 20, 1977, pp. 1385—1400.

Solution to the Backward-Kolmogorov Equation for a Nonstationary Oscillation Problem

G. P. Solomos[1] and P-T. D. Spanos[2]

The solution of a backward-Kolmogorov equation is presented. This equation is associated with a Markov approximation of the response amplitude of a lightly damped linear oscillator driven by an evolutionary random excitation.

Introduction

Uncertainties of loads on structures due to seismic motion, impact, strong gust, or sea waves can be adequately modeled by using nonstationary random processes. The theory for the determination of the exact response statistics of linear oscillators subjected to these kind of loads has been well developed. However, the exact solutions are frequently of limited practicality because of their mathematical complexity. Alternatively, approximate solutions that are mathematically more tractable can be sought. Furthermore, while the determination of the displacement and velocity statistics is the primary objective of such an analysis, the statistics of other response parameters, such as the response amplitude, can be equally important in certain cases [1-4]. The choice of this response parameter proves particularly convenient when some special conditions, considered in the next section, which regard the oscillator and the excitation, are satisfied.

Statement of the Problem

Consider the motion of an oscillator described by the equation

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = w(t), \tag{1}$$

where ζ and ω_0 are the damping ratio and the angular natural frequency, respectively. The excitation $w(t)$ is a nonstationary zero-mean random process possessing a broadband, evolutionary spectral density function $S(\omega,t)$ [5]. If the damping of the oscillator is light and the level of the excitation is low, it can be shown [4, 6] that the amplitude $a(t)$ of the response, defined by the equation

$$a(t) = \sqrt{x^2(t) + [\dot{x}(t)/\omega_0]^2}, \tag{2}$$

is approximately governed by the first-order stochastic differential equation

$$\dot{a}(t) = -\zeta\omega_0 a(t) + \frac{\pi S(\omega_0,t)}{2\omega_0^2 a(t)} + \frac{\sqrt{\pi S(\omega_0,t)}}{\omega_0}\eta(t), \tag{3}$$

where $\eta(t)$ is a stationary, zero-mean, white-noise process of unit intensity. Clearly, equation (3) indicates that $a(t)$ can be modeled as a one-dimensional Markov vector. As is known, of fundamental importance for the study of this type of process is its transition probability density function $q(a_1,t_1 \mid a,t)$, $t_1 > t$, defined as $q(a_1,t_1 \mid a,t)da_1 = \text{Prob.}[\text{the amplitude is between } a_1 \text{ and } a_1 + da_1 \text{ at time } t_1, \text{ given that it was between } a \text{ and } a+da \text{ at time } t]$. In this regard, statistical information about the amplitude can be obtained by solving the Fokker-Planck equation

$$\frac{\partial Q}{\partial t_1} = \frac{\partial}{\partial a_1}\left\{\left[\zeta\omega_0 a_1 - \frac{\pi S(\omega_0,t_1)}{2\omega_0^2 a_1}\right]Q\right\} + \frac{\pi S(\omega_0,t_1)}{2\omega_0^2}\frac{\partial^2 Q}{\partial a_1^2} \tag{4}$$

associated with equation (3). In equation (4) Q can be either the transition or the marginal density function of the amplitude. Based on equation (4) the marginal probability density function $p(a_1,t_1)$ of the amplitude, for zero initial conditions, has been determined in reference [4]. Specifically it has been found that

$$p(a_1,t_1) = \frac{a_1}{c_1}\exp\left(-\frac{a_1^2}{2c_1}\right), \tag{5}$$

where

$$c_1 \equiv c(t_1) = \frac{\pi}{\omega_0^2}\exp(-2\zeta\omega_0 t_1)\int_0^{t_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy. \tag{6}$$

Guided by this result and again through equation (4), the transition probability density function $q(a_1,t_1 \mid a,t)$, $t_1 > t$, has been determined in reference [7]. Specifically

$$q(a_1,t_1 \mid a,t) = \frac{a_1}{c(t,t_1)}\exp\left[-\frac{a_1^2 + a^2 e^{-2\zeta\omega_0(t_1-t)}}{2c(t,t_1)}\right] I_0\left(\frac{a_1 a e^{-\zeta\omega_0(t_1-t)}}{c(t,t_1)},\right) \tag{7}$$

where

$$c(t,t_1) = \frac{\pi}{\omega_0^2}\exp(-2\zeta\omega_0 t_1)\int_t^{t_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy, \tag{8}$$

and $I_0(\cdot)$ stands for the modified Bessel function of the first kind and of zero order. By using equation (7) the marginal probabilities and the moments of the amplitude, for fairly general initial conditions of the oscillator, can be obtained [7].

[1] Assistant Instructor of Engineering Mechanics, The University of Texas at Austin, Austin, Texas 78712.

[2] Associate Professor of Engineering Mechanics, The University of Texas at Austin, Austin, Texas 78712. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, May 1981; final revision, May, 1982.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(b) In the steady state the locations and sizes of the separation bubbles for Re = 1130, 1250, 1355, $\delta = 1.58$ are compared with Vogel's and Ronnenberg's observations in Fig. 3. In the numerical calculations the lower separation point shifts slightly upward when the mesh is refined. Ronnenberg's laser measurements are more accurate than those in Vogel's experiments. Hence, Fig. 3 shows that the agreement between computations and experiments is satisfactory.

Good agreement is obtained for the location of the limiting curve as shown in Fig. 1. Two regions of Vogel's curve were verified as follows: the computed flow fields for points in Fig. 1 outside the region of bubble occurrence exhibit no stagnation point; inside the region tiny bubbles appear, whose resolution is a function of the grid.

References

1 Peckham, D. H., and Atkinson, S. A., *British ARC CP*, No. 508, 1957.

2 Lugt, H. J., Ph.D. Thesis, University of Stuttgart, Germany, 1959 (short version translated by "The British Hydromechanics Research Ass." Report T 716, Feb. 1962).

3 Sarpkaya, T., "Vortex Breakdown in Swirling Conicad flows," *AIAA Journal*, Vol. 9, 1971, pp. 1792-1799.

4 Hall, M. G., *Annual Review of Fluid Mechanics* (Annual Review Inc., Palo Alto, Calif.) Vol. 4, 1972, pp. 195—218.

5 Leibovich, S., *Annual Review of Fluid Mechanics* (Annual Review Inc., Palo Alto, Calif.) Vol. 10, 1978, pp. 221—246.

6 Escudier, M. P., and Zehnder, N., "Vortex-Flow regimes," *J. of Fluid Mech.*, Vol. 115, 1982, pp. 105—121.

7 Lugt, H. J., and Haussling, H. J., *IUTAM Symposium on Unsteady Boundary Layers*, Les Presses de l'Universite Laval, Quebec, 1972, pp. 1366—1396.

8 Lugt, H. J., and Haussling, H. J., *Acta Mechanica*, Vol. 18, 1973, pp. 255—272.

9 Vogel, H. U., *Max-Planck-Institut für Strömungsforschung*, Gottingen, Bericht 6, 1968, pp. 1-91.

10 Ronnenberg, B., *Max-Planck-Institut für Stömungsforchung*, Gottingen, Bericht 20, 1977, pp. 1-73.

11 Grabowski, W. J., and Berger, S. A., *J. of Fluid Mech.*, Vol. 75, 1976, pp. 525—544.

12 Faler, J. H., and Leibovich, S., *Phys. of Fluids*, Vol. 20, 1977, pp. 1385—1400.

Solution to the Backward-Kolmogorov Equation for a Nonstationary Oscillation Problem

G. P. Solomos[1] and P-T. D. Spanos[2]

The solution of a backward-Kolmogorov equation is presented. This equation is associated with a Markov approximation of the response amplitude of a lightly damped linear oscillator driven by an evolutionary random excitation.

Introduction

Uncertainties of loads on structures due to seismic motion, impact, strong gust, or sea waves can be adequately modeled by using nonstationary random processes. The theory for the determination of the exact response statistics of linear oscillators subjected to these kind of loads has been well developed. However, the exact solutions are frequently of limited practicality because of their mathematical complexity. Alternatively, approximate solutions that are mathematically more tractable can be sought. Furthermore, while the determination of the displacement and velocity statistics is the primary objective of such an analysis, the statistics of other response parameters, such as the response amplitude, can be equally important in certain cases [1-4]. The choice of this response parameter proves particularly convenient when some special conditions, considered in the next section, which regard the oscillator and the excitation, are satisfied.

Statement of the Problem

Consider the motion of an oscillator described by the equation

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = w(t), \tag{1}$$

where ζ and ω_0 are the damping ratio and the angular natural frequency, respectively. The excitation $w(t)$ is a nonstationary zero-mean random process possessing a broadband, evolutionary spectral density function $S(\omega,t)$ [5]. If the damping of the oscillator is light and the level of the excitation is low, it can be shown [4, 6] that the amplitude $a(t)$ of the response, defined by the equation

$$a(t) = \sqrt{x^2(t) + [\dot{x}(t)/\omega_0]^2}, \tag{2}$$

is approximately governed by the first-order stochastic differential equation

$$\dot{a}(t) = -\zeta\omega_0 a(t) + \frac{\pi S(\omega_0,t)}{2\omega_0^2 a(t)} + \frac{\sqrt{\pi S(\omega_0,t)}}{\omega_0}\eta(t), \tag{3}$$

where $\eta(t)$ is a stationary, zero-mean, white-noise process of unit intensity. Clearly, equation (3) indicates that $a(t)$ can be modeled as a one-dimensional Markov vector. As is known, of fundamental importance for the study of this type of process is its transition probability density function $q(a_1,t_1 \mid a,t)$, $t_1 > t$, defined as $q(a_1,t_1 \mid a,t)da_1 = \text{Prob.}[\text{the amplitude is between } a_1 \text{ and } a_1 + da_1 \text{ at time } t_1, \text{ given that it was between } a \text{ and } a+da \text{ at time } t]$. In this regard, statistical information about the amplitude can be obtained by solving the Fokker-Planck equation

$$\frac{\partial Q}{\partial t_1} = \frac{\partial}{\partial a_1}\left\{\left[\zeta\omega_0 a_1 - \frac{\pi S(\omega_0,t_1)}{2\omega_0^2 a_1}\right]Q\right\} + \frac{\pi S(\omega_0,t_1)}{2\omega_0^2}\frac{\partial^2 Q}{\partial a_1^2} \tag{4}$$

associated with equation (3). In equation (4) Q can be either the transition or the marginal density function of the amplitude. Based on equation (4) the marginal probability density function $p(a_1,t_1)$ of the amplitude, for zero initial conditions, has been determined in reference [4]. Specifically it has been found that

$$p(a_1,t_1) = \frac{a_1}{c_1}\exp\left(-\frac{a_1^2}{2c_1}\right), \tag{5}$$

where

$$c_1 \equiv c(t_1) = \frac{\pi}{\omega_0^2}\exp(-2\zeta\omega_0 t_1)\int_0^{t_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy. \tag{6}$$

Guided by this result and again through equation (4), the transition probability density function $q(a_1,t_1 \mid a,t)$, $t_1 > t$, has been determined in reference [7]. Specifically

$$q(a_1,t_1 \mid a,t) = \frac{a_1}{c(t,t_1)}\exp\left[-\frac{a_1^2 + a^2 e^{-2\zeta\omega_0(t_1-t)}}{2c(t,t_1)}\right] I_0\left(\frac{a_1 a e^{-\zeta\omega_0(t_1-t)}}{c(t,t_1)},\right) \tag{7}$$

where

$$c(t,t_1) = \frac{\pi}{\omega_0^2}\exp(-2\zeta\omega_0 t_1)\int_t^{t_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy, \tag{8}$$

and $I_0(\cdot)$ stands for the modified Bessel function of the first kind and of zero order. By using equation (7) the marginal probabilities and the moments of the amplitude, for fairly general initial conditions of the oscillator, can be obtained [7].

[1] Assistant Instructor of Engineering Mechanics, The University of Texas at Austin, Austin, Texas 78712.

[2] Associate Professor of Engineering Mechanics, The University of Texas at Austin, Austin, Texas 78712. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, May 1981; final revision, May, 1982.

Copyright © 1982 by ASME

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

It is noted that equation (7) has been derived by guessing intuitively a solution form and determining the unknown functions involved in it.

The main objective of the present Brief Note is the presentation of an alternative direct procedure for the determination of $q(a_1,t_1\,|\,a,t)$. While it is possible to work with equation (4), its adjoint equation is employed instead, since the mathematical manipulations involved are substantially less in this case.

Method of Solution

The backward-Kolmogorov equation associated with equation (3) is

$$\frac{\partial q}{\partial t}=\left[\zeta\omega_0 a-\frac{\pi S(\omega_0,t)}{2\omega_0^2 a}\right]\frac{\partial q}{\partial a}-\frac{\pi S(\omega_0,t)}{2\omega_0^2}\frac{\partial^2 q}{\partial a^2}. \tag{9}$$

It is noted that this is a parabolic differential equation and that the domain of the amplitude is the set $a\geq 0$. Appended to equation (9) is the initial condition

$$q(a_1,t_1\,|\,a,t_1)=\delta(a-a_1), \tag{10}$$

where $\delta(\cdot)$ is the Dirac delta function, and the boundary conditions

$$q(a_1,t_1\,|\,\infty,t)=0, \tag{11}$$

$$q(a_1,t_1\,|\,0,t)=\text{finite}. \tag{12}$$

To recast equation (9) in a more tractable form, the following transformation of variables is introduced

$$z=\frac{a^2}{2c}, \tag{13}$$

$$s=t, \tag{14}$$

where for convenience symbols such as $a\equiv a(t)$, $a_1\equiv a(t_1), c\equiv c(0,t)=c(0,s), c_1\equiv c(0,t_1)=c(0,s_1)$, etc., will be used from this point on. In terms of the new variables the conditions expressed by equation (10) through equation (12) can be written as follows

$$q(z_1,s_1\,|\,z,s_1)=\frac{\sqrt{2z_1c_1}}{c}\,\delta\left(z-\frac{z_1c_1}{c}\right), \tag{15}$$

$$q(z_1,s_1\,|\,\infty,s)=0, \tag{16}$$

$$q(z_1,s_1\,|\,0,s)=\text{finite}. \tag{17}$$

After some chain-rule differentiations and rearrangement of terms equation (9) reduces to

$$\frac{\omega_0^2 c}{\pi S(\omega_0,s)}\frac{\partial q}{\partial s}=(z-1)\frac{\partial q}{\partial z}-z\frac{\partial^2 q}{\partial z^2}. \tag{18}$$

Following the usual separation of variables procedure, the solution of equation (18) is sought in the form

$$q=Z(z)\,T(s). \tag{19}$$

Substituting equation (19) into equation (18) leads to the equations

$$\frac{dT}{ds}=\lambda\frac{\pi S(\omega_0,s)}{\omega_0^2 c(0,s)}T, \tag{20}$$

$$z\frac{d^2Z}{dz^2}+(1-z)\frac{dZ}{dz}+\lambda Z=0, \tag{21}$$

where λ is a separation parameter.

To solve equation (20) observe that it can be written as

$$\frac{dT}{T}=\lambda\frac{d\left\{\int_0^S\exp(2\zeta\omega_0 y)S(\omega_0,y)dy\right\}}{\left\{\int_0^S\exp(2\zeta\omega_0 y)S(\omega_0,y)dy\right\}}, \tag{22}$$

and therefore

$$T(s)=\left[\int_0^S\exp(2\zeta\omega_0 y)S(\omega_0,y)dy/k_1(s_1)\right]^{\lambda}, \tag{23}$$

where $k_1(s_1)$ is a function of s_1. For reasons to be seen subsequently the function $k_1(s_1)$ is chosen such that the quantity $[\int_0^S\exp(2\zeta\omega_0 y)\,S(\omega_0,y)dy/k_1(s_1)]$ is always less than one for $s<s_1$, and equal to one when $s=s_1$. The requirement is satisfied by taking

$$T(s)=\left[\int_0^S\exp(2\zeta\omega_0 y)S(\omega_0,y)dy/\int_0^{s_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy\right]^{\lambda}=r^{2\lambda}, \tag{24}$$

where

$$r^2=\frac{\int_0^S\exp(2\zeta\omega_0 y)S(\omega_0,y)dy}{\int_0^{s_1}\exp(2\zeta\omega_0 y)S(\omega_0,y)dy}=\frac{c}{c_1}e^{-2\zeta\omega_0(t_1-t)}. \tag{25}$$

Interestingly, it can be proved that the parameter r^2 is the correlation coefficient of the process $a^2/2c$.

It is clear that equation (21) is a confluent hypergeometric equation. The solution of this equation, finite at $z=0$, can be expressed in terms of the confluent hypergeometric function $M(\cdot)$ [8], in the form

$$Z(z)=M(-\lambda,1,z). \tag{26}$$

Satisfaction of the boundary condition (16) requires that

$$Z(\infty)=M(-\lambda,1,\infty)=0, \tag{27}$$

which is an "algebraic equation" yielding for λ the values

$$\lambda_n=n,\quad n=0,1,2,\ldots \tag{28}$$

Thus,

$$M(-n,1,z)=L_n(z), \tag{29}$$

where $L_n(z)$ is the Laguerre polynomial of order n. Furthermore

$$\int_0^{\infty}L_n(z)L_l(z)e^{-z}dz=0;\quad n,l=\text{integers},n\neq l \tag{30}$$

Based on equations (24), (28), and (29) the solution of equation (18) is therefore constructed as

$$q(z_1,s_1\,|\,z,s)=\sum_{n=0}^{\infty}H_n L_n(z)r^{2n} \tag{31}$$

with H_n being functions of z_1 and s_1 to be determined by using equation (15). In this regard, it is observed that when $t\to t_1$ then $c\to c_1$ and $r^2\to 1$. Thus,

$$\frac{\sqrt{2z_1c_1}}{c_1}\,\delta\left(z-\frac{z_1c_1}{c_1}\right)=\sum_{n=0}^{\infty}H_n L_n(z). \tag{32}$$

Multiplying equation (32) by $e^{-z}L_i(z)$, integrating over z from 0 to ∞, and taking into account equation (30), yields

$$H_i=\sqrt{\frac{2z_1}{c_1}}e^{-z_1}L_i(z_1). \tag{33}$$

Equation (31) is then rewritten as

$$q(z_1,s_1\,|\,z,s)=\sqrt{\frac{2z_1}{c_1}}e^{-z_1}\sum_{n=0}^{\infty}L_n(z)L_n(z_1)r^{2n}, \tag{34}$$

and changing back to the old variables gives

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$q(a_1,t_1\,|\,a,t)=\frac{a_1}{c_1}\exp\left(-\frac{a_1^2}{2c_1}\right)$$

$$\sum_{n=0}^{\infty}L_n\left(\frac{a^2}{2c}\right)L_n\left(\frac{a_1^2}{2c_1}\right)r^{2n}. \tag{35}$$

A last step that transforms equation (35) into equation (7) includes utilization of the Hille-Hardy generating function formula [8]. Specifically, equation (35) becomes

$$q(a_1,t_1\,|\,a,t)=\frac{a_1}{c_1(1-r^2)}\exp\left[-\frac{a_1^2c+a^2c_1r^2}{2c_1c(1-r^2)}\right]$$

$$I_0\left(\frac{raa_1}{\sqrt{c_1c}(1-r^2)}\right). \tag{36}$$

It can be readily verified by substituting the expressions for c and c_1, that equation (36) is identical with equation (7).

Summary

The transition probability density function $q(a_1,t_1\,|\,a,t)$ of a Markovian approximation of the response amplitude of an oscillator under nonstationary excitation has been determined in an analytical form. Toward this end, the associated backward-Kolmogorov equation has been solved by transforming it to a form amenable to solution by the method of separation of variables. This procedure has led to a boundary value problem which has been solved by using an infinite series of Laguerre polynomials. The infinite series solution is equivalent to a closed-form solution involving a Bessel function.

Acknowledgments

Financial support of this study by the Grant NSF-PFR8006569 from the National Science Foundation and the Grant NAG 3-210 from the NASA, Lewis Research Center are gratefully acknowledged.

References

1 Yang, J. N., "Non-Stationary Envelope Process and First Excursion Probability," *Journal of Structural Mechanics*, Vol. 1, 1972, pp. 231–248.

2 Ariaratnam, S. T., and Pi, H. N., "On the First-Passage Time for Envelope Crossing for a Linear Oscillator," *International Journal of Control*, Vol. 18, 1973, pp. 89–96.

3 Lennox, W. C., and Fraser, D. A., "On the First-Passage Distribution for the Envelope of a Non-Stationary Narrow-Band Stochastic Process," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 793–797.

4 Spanos, P-T. D., and Lutes, L. D., "Probability of Response to Evolutionary Process," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No, EM2, Proc. Paper 15313, Apr. 1980, pp. 213–224.

5 Shinozuka, M., "Random Processes With Evolutionary Power," *Journal of the Engineering Mechanics Division*, ASCE, No. EM4, Vol. 96, Proc. Paper 7444, Aug. 1970, pp. 543–545.

6 Spanos, P-T. D., "Probabilistic Earthquake Energy Spectra Equations," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No. EM1, Proc. Paper 15204, Feb. 1980, pp. 147–159.

7 Spanos, P-T. D., and Solomos, G. P., "On the Statistics of the Amplitude of Oscillator Response to Random Excitation With Evolutionary Power Spectrum," TICOM Report 81-2, May 1981, The University of Texas at Austin.

8 Erdelyi, A. et al., *Higher Transcendental Functions*, McGraw-Hill, New York, 1953.

A Note on the Length and Displacements of a Dugdale or Craze Zone

T. J. Lardner[1]

[1] School of Engineering, Department of Civil Engineering, University of Massachusetts, Amherst, Mass. 01003. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, February, 1982; final revision, May, 1982.

Introduction

The analysis of the stress distribution in an infinite two-dimensional, isotropic, linear-elastic material containing a central crack with an extended Dugdale Zone is well known; see, for example [1–9] and the references therein. The length of the Dugdale Zone is determined from the requirement that the normal stress at the tip of the Dugdale Zone is non-singular. This requirement is met by canceling the stress intensity factor from the uniform stress with the stress intensity factor from the stress distribution in the Dugdale Zone. These concepts have been carried over to the analysis of crazes ahead of cracks in polymeric materials [3–9].

An important problem that arises in the stress analysis of polymers is the effect of material properties on the distribution of stresses about a crack and craze, e.g., in an oriented polymer, [4, p. 103]. A related question is the effect orthotropic material properties may have on the length of the Dugdale Zone.

It turns out that the orthotropic stress-strain material constants have no effect on the length of the Dugdale Zone for symmetric deformations. This result follows from the observation that the stress intensity factor in an orthotropic material for symmetric deformations (mode I deformation) is identical to the stress intensity factor in an isotropic material, [10–13]. Since the length of the Dugdale Zone is found by elimination of the stress singularity, the material constants have no effect on the length.

However, the direct application of the stress intensity factor approach may lead to difficulties ([3]; also see comment in [4], p. 117). Further, if the cohesive stress in the Dugdale Zone is variable, an additional calculation is necessary; most often the approach to this calculation is by complex variable methods.

The purpose of this Brief Note is to highlight some simple results for crack and craze problems through the use of integral transform methods, reference [14], in contrast to the usual derivation of the results using complex variable methods.

Formulation

The stress strain relations for an orthotropic material can be written in the form

$$E_ye_x=\beta^2\sigma_x-\nu_y\sigma_y;\quad E_ye_y=\sigma_y-\nu_y\sigma_x$$

$$E_ye_{xy}=(E_y/2G)\sigma_{xy};\quad \beta^2=(E_y/E_x)=(\nu_y/\nu_x)$$

The crack and craze zone are assumed to lie on the x axis; $\hat{c}$ is the length of the crack and $\hat{a}$ is the length of crack and craze zone. The normal stress on the crack is the stress σ_0 at infinity, and the stress distribution in the craze zone is a known distribution. The length of the craze zone will depend on the resultant of the stress distribution in the craze zone.

We will follow closely the approach in [14] but we do not give the details here. We nondimensionalize the x coordinate with respect to the total length of the crack and craze. Upon use of Fourier transforms and the assumption of symmetric deformation about the y axis, we find that for orthotropic materials the problem reduces to the same pair of dual integral equations in [14], equations (2.36) and (2.38), for isotropic materials. This result follows from the two-dimensional nature of the problem and from the homogeneous boundary conditions on the shear stress and the vertical displacement.

The normal stress ahead of the craze, [14], is

$$\sigma_y(x,0)=\frac{2}{\pi}\left[\frac{x}{\sqrt{x^2-1}}q(1)+x\int_0^1\frac{q'(t)}{\sqrt{x^2-t^2}}dt\right]$$

where

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$q(a_1,t_1\,|\,a,t)=\frac{a_1}{c_1}\exp\left(-\frac{a_1^2}{2c_1}\right)$$

$$\sum_{n=0}^{\infty}L_n\left(\frac{a^2}{2c}\right)L_n\left(\frac{a_1^2}{2c_1}\right)r^{2n}. \tag{35}$$

A last step that transforms equation (35) into equation (7) includes utilization of the Hille-Hardy generating function formula [8]. Specifically, equation (35) becomes

$$q(a_1,t_1\,|\,a,t)=\frac{a_1}{c_1(1-r^2)}\exp\left[-\frac{a_1^2c+a^2c_1r^2}{2c_1c(1-r^2)}\right]$$

$$I_0\left(\frac{raa_1}{\sqrt{c_1}c(1-r^2)}\right). \tag{36}$$

It can be readily verified by substituting the expressions for c and c_1, that equation (36) is identical with equation (7).

Summary

The transition probability density function $q(a_1,t_1\,|\,a,t)$ of a Markovian approximation of the response amplitude of an oscillator under nonstationary excitation has been determined in an analytical form. Toward this end, the associated backward-Kolmogorov equation has been solved by transforming it to a form amenable to solution by the method of separation of variables. This procedure has led to a boundary value problem which has been solved by using an infinite series of Laguerre polynomials. The infinite series solution is equivalent to a closed-form solution involving a Bessel function.

Acknowledgments

Financial support of this study by the Grant NSF-PFR8006569 from the National Science Foundation and the Grant NAG 3-210 from the NASA, Lewis Research Center are gratefully acknowledged.

References

1 Yang, J. N., "Non-Stationary Envelope Process and First Excursion Probability," *Journal of Structural Mechanics*, Vol. 1, 1972, pp. 231–248.

2 Ariaratnam, S. T., and Pi, H. N., "On the First-Passage Time for Envelope Crossing for a Linear Oscillator," *International Journal of Control*, Vol. 18, 1973, pp. 89–96.

3 Lennox, W. C., and Fraser, D. A., "On the First-Passage Distribution for the Envelope of a Non-Stationary Narrow-Band Stochastic Process," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, 1974, pp. 793–797.

4 Spanos, P-T. D., and Lutes, L. D., "Probability of Response to Evolutionary Process," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No, EM2, Proc. Paper 15313, Apr. 1980, pp. 213–224.

5 Shinozuka, M., "Random Processes With Evolutionary Power," *Journal of the Engineering Mechanics Division*, ASCE, No. EM4, Vol. 96, Proc. Paper 7444, Aug. 1970, pp. 543–545.

6 Spanos, P-T. D., "Probabilistic Earthquake Energy Spectra Equations," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No. EM1, Proc. Paper 15204, Feb. 1980, pp. 147–159.

7 Spanos, P-T. D., and Solomos, G. P., "On the Statistics of the Amplitude of Oscillator Response to Random Excitation With Evolutionary Power Spectrum," TICOM Report 81-2, May 1981, The University of Texas at Austin.

8 Erdelyi, A. et al., *Higher Transcendental Functions*, McGraw-Hill, New York, 1953.

A Note on the Length and Displacements of a Dugdale or Craze Zone

T. J. Lardner[1]

[1] School of Engineering, Department of Civil Engineering, University of Massachusetts, Amherst, Mass. 01003. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, February, 1982; final revision, May, 1982.

Introduction

The analysis of the stress distribution in an infinite two-dimensional, isotropic, linear-elastic material containing a central crack with an extended Dugdale Zone is well known; see, for example [1–9] and the references therein. The length of the Dugdale Zone is determined from the requirement that the normal stress at the tip of the Dugdale Zone is non-singular. This requirement is met by canceling the stress intensity factor from the uniform stress with the stress intensity factor from the stress distribution in the Dugdale Zone. These concepts have been carried over to the analysis of crazes ahead of cracks in polymeric materials [3–9].

An important problem that arises in the stress analysis of polymers is the effect of material properties on the distribution of stresses about a crack and craze, e.g., in an oriented polymer, [4, p. 103]. A related question is the effect orthotropic material properties may have on the length of the Dugdale Zone.

It turns out that the orthotropic stress-strain material constants have no effect on the length of the Dugdale Zone for symmetric deformations. This result follows from the observation that the stress intensity factor in an orthotropic material for symmetric deformations (mode I deformation) is identical to the stress intensity factor in an isotropic material, [10–13]. Since the length of the Dugdale Zone is found by elimination of the stress singularity, the material constants have no effect on the length.

However, the direct application of the stress intensity factor approach may lead to difficulties ([3]; also see comment in [4], p. 117). Further, if the cohesive stress in the Dugdale Zone is variable, an additional calculation is necessary; most often the approach to this calculation is by complex variable methods.

The purpose of this Brief Note is to highlight some simple results for crack and craze problems through the use of integral transform methods, reference [14], in contrast to the usual derivation of the results using complex variable methods.

Formulation

The stress strain relations for an orthotropic material can be written in the form

$$E_ye_x=\beta^2\sigma_x-\nu_y\sigma_y;\quad E_ye_y=\sigma_y-\nu_y\sigma_x$$

$$E_ye_{xy}=(E_y/2G)\sigma_{xy};\quad \beta^2=(E_y/E_x)=(\nu_y/\nu_x)$$

The crack and craze zone are assumed to lie on the x axis; $\hat{c}$ is the length of the crack and $\hat{a}$ is the length of crack and craze zone. The normal stress on the crack is the stress σ_0 at infinity, and the stress distribution in the craze zone is a known distribution. The length of the craze zone will depend on the resultant of the stress distribution in the craze zone.

We will follow closely the approach in [14] but we do not give the details here. We nondimensionalize the x coordinate with respect to the total length of the crack and craze. Upon use of Fourier transforms and the assumption of symmetric deformation about the y axis, we find that for orthotropic materials the problem reduces to the same pair of dual integral equations in [14], equations (2.36) and (2.38), for isotropic materials. This result follows from the two-dimensional nature of the problem and from the homogeneous boundary conditions on the shear stress and the vertical displacement.

The normal stress ahead of the craze, [14], is

$$\sigma_y(x,0)=\frac{2}{\pi}\left[\frac{x}{\sqrt{x^2-1}}q(1)+x\int_0^1\frac{q'(t)}{\sqrt{x^2-t^2}}dt\right]$$

where

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$q(t)=\int_0^t \frac{p(u)}{\sqrt{t^2-u^2}}\,du$$

and $p(u)$ is the known normal stress distribution for $x<1$.

It immediately follows that to eliminate the singularity at $x=1$, corresponding to the tip of the craze, we should set

$$q(1)=0 \tag{1}$$

Equation (1) is the equation for the determination of the length of the craze zone and is equivalent to the result given, for example, by Kramer, [4, equation (19*b*)].

If the stress in the craze zone is constant so that $p=(\sigma_0-\sigma_1)$, $c<u<1$, where σ_0 is the applied stress and σ_1 is the craze stress, equation (1) gives the well-known result for the length of the craze zone upon integration and rewriting the equation in the original variables; see [4, equation (19)].

If, as was discussed in [5], (see also [4] and [9]) we consider in addition to the crack and craze, a plastic zone in front of the craze, the condition given by equation (1) still applies. This equation to determine the length of the zone now becomes (where we now use the notation of [4, Fig. 8]):

$$\sigma_o\int_0^c \frac{du}{\sqrt{1-u^2}}+\int_c^{1-r}\frac{(\sigma_0-S_c)}{\sqrt{1-u^2}}\,du + \int_{1-r}^{r}\frac{(\sigma_0-S_t)}{\sqrt{1-u^2}}\,du=0$$

We see that in this form the inclusion of a variable stress in the craze and plastic zone can be effected easily. For constant S_c and S_t the equation for the determination of $\hat{a}$ follows upon integration. The result agrees with [4], equation (28); (also [7], equation (37)). If the strength of the singularity at the craze is to be selected as some equilibrium value or initiation value for the material, then a modification of equation (1) can be used to specify the strength of the singularity.

For the more general case of a stepwise distribution of stress in the craze zone, $p(u)=\sigma_k-\sigma_0$, $c_{k-1}<x<c_k$, the expression for $q(t)$ can be written in the form

$$q(t)=-\sigma_0\frac{\pi}{2}+\sigma_1\cos^{-1}(c_0/t)H(t-c_0) + \sum_{k=1}^{N-1}(\sigma_{k+1}-\sigma_k)\cos^{-1}(c_k/t)H(t-c_k)$$

where $H(x)$ is the Heaviside step function, c_0 is the crack tip, and $c_n=1$.

The expression for the normal displacement can be written in the form (see [14], equation (2.3.13) for the isotropic case)

$$w=2w_0\int_x^1 \frac{tq(t)}{\sqrt{t^2-x^2}}\,dt;\quad w_0=[(E_y/G)-2\nu_y+2\beta]^{1/2}/E_y$$

Upon substitution for $q(t)$ and integration by parts, we find

$$\frac{w\pi}{2w_0\hat{a}}=\sqrt{1-x^2}\,q(1)+\sigma_1 c_0\int_{\max(x,c_0)}^1\sqrt{\frac{t^2-x^2}{t^2-c_0^2}}\frac{dt}{t} - \sum_{k=1}^{N-1}(\sigma_{k+1}-\sigma_k)c_k\int_{\max(x,C_0)}^1\sqrt{\frac{t^2-x^2}{t^2-c_k^2}}\frac{dt}{t}$$

In the special case when $N=1$ corresponding to a constant stress σ_1 in the craze zone, we have in terms of the dimensional quantities,

$$w=\frac{2w_0\sigma_1\hat{a}c}{\pi}\int_{\max(x,c)}^1\sqrt{\frac{t^2-x^2}{t^2-c^2}}\frac{dt}{t}=\frac{w_0\sigma_1\hat{a}}{2\pi}H(x;c)$$

The expression for H from this equation is an alternative expression to the closed-form expression for H given in [4, equation (18)].

In the case of an isotropic material

$$w=\frac{\sigma_1\hat{a}}{\pi E}H(x;c)$$

is in agreement with [4]. We note that the displacement in the crack and craze zone for an orthotropic material is a numerical factor times the displacement for an isotropic material, namely, (see also [13], equation (10)),

$$w_{\text{orth}}=(w_0E/2)w_{\text{iso}}$$

The crack opening displacement follows as

$$w=\frac{w_0\sigma_1\hat{a}}{2\pi}(-4c\ln c)$$

Expansions of the integral around $x=0$, c, and 1 can easily be carried out to obtain approximate analytical expressions for the displacements in the crack and craze zone.

Conclusion

We have presented a number of observations and results for the stress and displacement calculations of symmetric craze zones. We noted that there is no effect of orthotropic material constants on the length of the craze zone, and we obtained a simple formula, equation (1), to be used to find the length of the zone. In addition, we derived an alternative expression for the displacement in the crack and craze zone.

References

1 Goodier, J., and Field, F., "Plastic Energy Dissipation in Crack Propagation," *Fracture of Solids 103–118*, Vol. 20 in *Metallurgical Soc. Confs.*, Drucker, D., and Gilman, J., eds., Gordon and Breach, N.Y., 1963.

2 Rice, J. R., "Mathematical Analysis in the Mechanics of Fracture," *Fracture, Vol. II*, Academic Press, N.Y., 1968, Chapter 3, Section C.2.

3 Gerberich, W., "A Dugdale—Baren-Blatt Model for an Equilibrium Craze," *Int. J. of Fracture*, Vol. 13, 1977, pp. 535–538.

4 Kramer, E. J., "Environmental Cracking of Polymers," in *Developments in Polymer Fracture*, Chapter 3, Andrews, E. H., ed., Applied Science Publishers, 1978.

5 Verheulpen-Heymans, N., and Bauwens, J. C., "Effect of Stress and Temperature on Dry Craze Growth Kinetics During Low-Stress Creep of Polycarbonate, Part 2, Theoretical Model," *J. Mat. Sci.*, Vol. 11, 1976, pp. 7–16.

6 Kramer, E. J., Krentz, H. G., and Ast, D. G., "Growth Criteria for Solvent Crazes," *J. Mat. Sci.*, Vol. 13, 1978, pp. 1093–1098.

7 Israel, S. J., Thomas, E. L., and Gerberich, W. W., "Correlations of the Craze Profile in PMMA with Dugdale's Plastic Zone Profile," *J. Mat. Sci.*, Vol. 14, 1978, pp. 2128–2138.

8 Doll, W., Seidelmann, V., and Konczol, L., "On the Validity of the Dugdale Model for Craze Zones at Crack Tips in PMMA," *J. Mat. Sci.*, Vol. 15, 1980, pp. 2389–2394.

9 Kramer, E. J., and Bubeck, R. A., "Growth Kinetics of Solvent Crazes in Glassy Polymers," *J. Polymer Science*, Vol. 16, 1978, pp. 1195–1217.

10 Sih, G. C., Paris, P. C., and Irwin, G. R., "On Cracks in Rectilinearly Anisotropic Bodies," *Int. J. of Fracture*, Vol. 1, 1965, pp. 189–203.

11 Ang, D. D., and Williams, M. L., "Combined Stresses in an Orthotropic Plate Having a Finite Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, 1961, pp. 372–378.

12 Bowie, O. L., "Solutions of Plane Crack Problems by Mapping Techniques," *Methods of Analysis and Solutions of Crack Problems*, Sih, G. C., ed., Noordhoff, 1973, Chapter 1.

13 Mendelson, A., and Sero, A., "Elastic Stress Distribution in a Finite-Width Orthotropic Plate Containing a Crack," NASA Report E-1576, 1964.

14 Sneddon, I. N., and Lowengrub, M., "Crack Problems in the Classical Theory of Elasticity," *Siam Series in Applied Math.*, Wiley, New York, 1969, Section 2.3.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Journal of Applied Mechanics

Discussion

Optimization in a Class of Hole Shapes in Plate Structures[1]

K. Rajaiah[2] and N. K. Naik[2]. The determination of optimum hole shapes to achieve minimum stress concentration is an important engineering problem. The author's analytical contribution is a welcome addition to this area. In this context, the discussers would like to draw attention to their recent work on optimum hole-shape determination in infinite orthotropic plates under inplane loading by conformal transformation [1]. An extension of this approach to quasi-rectangular holes has also been reported separately [2]. In this approach, following the Lekhnitskii's method of analysis [3], the hole-shape parameter ϵ and hence the hole shape itself was varied until the maximum tangential stress on the hole boundary became a minimum. For this optimum condition, it is found that the tangential stress distribution becomes nearly uniform on the tensile segment of the boundary. A comparison of the results for isotropic plates reported in reference [1] with those given by Dhir, is given in Table 1, which shows very close correlation. Table 1 also contains additional information for the case of orthotropic plates.

An interesting point to observe is the stress distribution achieved which is not absolutely uniform even along the tensile segment of the hole boundary in Dhir's as well as our work [1] for the case of uniaxial loading condition. On the compressive segment, the distribution is still worse and minimum stress values are not achieved. Whereas, with our approach [1], one can independently minimize the stress concentrations on the tensile and compressive segments of the hole boundary by choosing different optimum values of ϵ for the two segments, this does not appear to be possible by Dhir's approach even though he has followed a minimization principle. Clearly, when the applied load becomes compressive (reversal of loading), the importance of the respective segments gets reversed and in that case, the numerically smaller tensile stress becomes important from a design point of view. Perhaps another possibility would be to choose a superior transformation function to optimize both the segments simultaneously. However as mentioned by Dhir in his Example II, there can be considerable complexity in the algebraic manipulation in such a case. With the experimental approach, one has the advantage of obtaining real optimized shape on both the segments and dealing with finite size plates [4], although one has to conduct a series of experiments to cover a large range of parameters.

[1] By S. K. Dhir, and published in the December, 1981, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 905-908.

[2] Professor and Lecturer, respectively, Aeronautical Engineering Department, Indian Institute of Technology, Bombay-400 076, India.

Table 1

Loading	Stress concentration factor (s.c.f.)			
	Optimized hole		Square hole	Circular hole
	[1]	[Dhir]		
Isotropic				
Uniaxial	2.47	2.47	2.85	3.00
Biaxial $\sigma_0:\sigma_0$	2.00	2.00	4.00	2.00
Biaxial $\sigma_0:-\sigma_0$	3.06	3.07	3.10	4.00
Orthotropic				
Uniaxial $E_1/E_2=3.41$	3.02	—	3.21	4.60
Uniaxial $E_1/E_2=0.293$	2.38	—	2.68	2.95
Biaxial, $\sigma_0:\sigma_0$ $E_1/E_2=3.41$	2.10	—	3.37	2.75
Biaxial, $\sigma_0:-\sigma_0$ $E_1/E_2=3.41$	4.23	—	4.25	6.45

References

1 Rajaiah, K., and Naik, N. K., "Optimum Hole Shape Determination in Infinite Orthotropic Plates Under Inplane Loading by Conformal Transformation," to appear in *Mechanics Research Communications*.

2 Rajaiah, K., and Naik, N. K., "Quasi-Rectangular Holes for Minimizing Stress Concentrations in Infinite Plates Under Inplane Loading," Technical Report No. AE TR 811201, I.I.T. Bombay, India, 1981.

3 Lekhnitskii, S. G., *Anisotropic Plates*, Gordon and Breach, New York, 1968.

4 Durelli, A. J., and Rajaiah, K., "Optimum Hole Shapes in Finite Plates Under Uniaxial Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, 1979, pp. 691-695.

Author's Closure

The discussion by Rajaiah and Naik is very much appreciated. It is somwhat unfortunate that both the references [1] (not yet published) and [2] mentioned by Rajaiah and Naik in their discussion are presently not available to the author. As a result it is difficult to comment on the advantages and/or disadvantages of their approach. However, it seems that independent minimization of the stresses on tensile and compressive segments of the opening would not necessarily lead to a truly optimized hole; especially because both compressive and tensile segments of the opening are not considered simultaneously. As was suggested in the original paper, the use of Lagrange multipliers would perhaps also be a viable possibility to accomplish the simultaneous optimization of both segments of the opening. For the information of the discussers the author has also successfully

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

used his technique of minimizing a stress-energy density integral to optimize the material properties of an orthotropic plate with an opening [1].

Reference

1 Dhir, S. K., "Optimization of Openings in Plates," AIAA/ASME/ASCE/AHS 23rd SDM Conference, New Orleans, May 1982.

Edge Effects in Laminated Composite Plates[1]

C. O. Horgan.[2] The authors have presented an interesting finite element approach to the analysis of edge and interlaminar effects in laminated composites. Their results demonstrate that numerical methods involving self-equilibrated eigenfunctions for such problems are most efficiently carried out using finite elements.

However, it should be noted that a purely numerical scheme such as that given by the authors does not reveal the explicit dependence of the *rate of decay* of edge effects on material and geometric parameters. Such results can be obtained from the qualitative analyses of Saint-Venant's principle [1–6] cited by the authors. (See also [1, 2] of this discussion for a summary of results in the anisotropic case.) For example, for plane strain of an homogeneous rectangular strip, transversely isotropic about the axial direction, it is shown by I. Choi and the discusser in [5] that the exponential stress decay rate has the asymptotic characterization

$$\gamma \sim 2\pi (G_{LT}/E_L)^{1/2}/b \quad \text{as} \quad G_{LT}/E_L \to 0, \tag{1}$$

for highly anisotropic materials. Here b is the strip width and the remaining notation is that of the paper under discussion. For the graphite/epoxy composite considered by the authors and in [5] (for which $G_{LT}/E_L = 0.03$), equation (1) yields the result $\gamma \sim 1.088/b$, which is in excellent agreement with the exact value $\gamma = 1.128/b$. We advocate use of the formula (1) for design purposes for fiber-reinforced composites (in both plane strain and generalized plane stress). Another situation considered in [6] is concerned with plane strain of a sandwich strip, composed of identical homogeneous *isotropic* face materials occupying two layers of equal thickness enclosing a dissimilar homogeneous *isotropic* core. The result analogous to (1) has the form (see equation (4.1) of [6])

$$\gamma\,(\text{total strip width}) \sim 2\left[\frac{2(f^2-3f+3)(1-\nu_f^2)E_c}{f^3(1-f)(1+\nu_c)E_f}\right]^{1/2}, \quad \text{as} \quad E_c/E_f \to 0. \tag{2}$$

Here f denotes the volume fraction of face material and E, and ν are the respective Young's moduli and Poisson ratios. It would be of interest to extend (2) to the case of laminates with anisotropic layers, as treated by the authors.

Finally, we concur with the authors' concluding remarks regarding the need for consideration of stress singularities in any discussion of edge effects in laminated plates.

References

1 Horgan, C. O., "Saint-Venant's Principle in Anisotropic Elasticity Theory," in: *Mechanical Behavior of Anisotropic Solids*, J. P. Boehler, ed., Proc. of Euromech Colloquium 115, Villard de Lans, 1979; Editions Scientifiques du CNRS, Paris, France, 1982, pp. 855–870.

2 Horgan, C. O., "Saint-Venant End Effects in Composites," *Journal of Composite Materials*, Vol. 12, 1982, in press.

Authors' Closure

There is no question regarding the usefulness and desirability of formulas describing the fundamental exponential decay rate of self-equilibrated edge effects in terms of material and geometrical parameters. However, such formulas, in the authors' opinion, are possible for a limited number of composite plates only, such as those cited by the discusser and perhaps for cross-ply and angle-ply plates with equal thickness laminates. Beyond these, the large number of parameters that must be reflected in meaningful formulas will preclude their derivations, simply due to algebraic difficulties.

The numerical analysis outlined in our paper enables an intuitive understanding of a large class of composite plate to be easily gained. The relatively low computational effort must be emphasized. Also, a great deal of information is revealed in a straightforward manner, e.g., the fundamental decay rate as well as the higher order ones and their displacement and stress distributions over the plate's cross section. Although not as elegant as a formula, the numerical technique is equally effective in sorting out the governing parameters on decay rates.

It would be a nice outcome if the numerical studies could lead to relatively simple formulas containing the essential nature of the rate of exponential decay.

A Simple and Effective Pipe Elbow Element – Interaction Effects[1]

J. F. Whatham.[2] I would like to point out that while the experimental points in Figs. 9–13 were taken from Whatham [1] and suitably acknowledged, no mention was made of the analytic solution in that paper. The results from that solution were closer to the experimental points than were the curves presented, particularly for circumferential stress; references [2–5] describe the method more fully.

The Bathe-Almeida model with only two boundary conditions cannot adequately describe the true situation in a curved pipe at a rigid flange where four strain components must vanish.

A computer program package BENDPAC written in FORTRAN IV and ASSEMBLER for an IBM 3031 computer is available from the Australian Atomic Energy Commission,

[1] By S. B. Dong and D. B. Goetschel, and published in the March, 1982, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 129–135.

[2] Professor of Mechanics, Department of Metallurgy, Mechanics and Materials Science, Michigan State University, East Lansing, Mich. 48824.

[1] By K. J. Bathe and C. A. Almeida, and published in the March, 1982 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 165–171.

[2] Australian Atomic Energy Commission, Lucas Heights Research Laboratories, New Ittawarra Road, Lucas Heights, NSW Australia.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

used his technique of minimizing a stress-energy density integral to optimize the material properties of an orthotropic plate with an opening [1].

Reference

1 Dhir, S. K., "Optimization of Openings in Plates," AIAA/ASME/ASCE/AHS 23rd SDM Conference, New Orleans, May 1982.

Edge Effects in Laminated Composite Plates[1]

C. O. Horgan.[2] The authors have presented an interesting finite element approach to the analysis of edge and interlaminar effects in laminated composites. Their results demonstrate that numerical methods involving self-equilibrated eigenfunctions for such problems are most efficiently carried out using finite elements.

However, it should be noted that a purely numerical scheme such as that given by the authors does not reveal the explicit dependence of the *rate of decay* of edge effects on material and geometric parameters. Such results can be obtained from the qualitative analyses of Saint-Venant's principle [1–6] cited by the authors. (See also [1, 2] of this discussion for a summary of results in the anisotropic case.) For example, for plane strain of an homogeneous rectangular strip, transversely isotropic about the axial direction, it is shown by I. Choi and the discusser in [5] that the exponential stress decay rate has the asymptotic characterization

$$\gamma \sim 2\pi (G_{LT}/E_L)^{1/2}/b \quad \text{as} \quad G_{LT}/E_L \to 0, \tag{1}$$

for highly anisotropic materials. Here b is the strip width and the remaining notation is that of the paper under discussion. For the graphite/epoxy composite considered by the authors and in [5] (for which $G_{LT}/E_L = 0.03$), equation (1) yields the result $\gamma \sim 1.088/b$, which is in excellent agreement with the exact value $\gamma = 1.128/b$. We advocate use of the formula (1) for design purposes for fiber-reinforced composites (in both plane strain and generalized plane stress). Another situation considered in [6] is concerned with plane strain of a sandwich strip, composed of identical homogeneous *isotropic* face materials occupying two layers of equal thickness enclosing a dissimilar homogeneous *isotropic* core. The result analogous to (1) has the form (see equation (4.1) of [6])

$$\gamma\,(\text{total strip width}) \sim 2\left[\frac{2(f^2-3f+3)(1-\nu_f^2)E_c}{f^3(1-f)(1+\nu_c)E_f}\right]^{1/2}, \quad \text{as} \quad E_c/E_f \to 0. \tag{2}$$

Here f denotes the volume fraction of face material and E, and ν are the respective Young's moduli and Poisson ratios. It would be of interest to extend (2) to the case of laminates with anisotropic layers, as treated by the authors.

Finally, we concur with the authors' concluding remarks regarding the need for consideration of stress singularities in any discussion of edge effects in laminated plates.

References

1 Horgan, C. O., "Saint-Venant's Principle in Anisotropic Elasticity Theory," in: *Mechanical Behavior of Anisotropic Solids*, J. P. Boehler, ed., Proc. of Euromech Colloquium 115, Villard de Lans, 1979; Editions Scientifiques du CNRS, Paris, France, 1982, pp. 855–870.

2 Horgan, C. O., "Saint-Venant End Effects in Composites," *Journal of Composite Materials*, Vol. 12, 1982, in press.

Authors' Closure

There is no question regarding the usefulness and desirability of formulas describing the fundamental exponential decay rate of self-equilibrated edge effects in terms of material and geometrical parameters. However, such formulas, in the authors' opinion, are possible for a limited number of composite plates only, such as those cited by the discusser and perhaps for cross-ply and angle-ply plates with equal thickness laminates. Beyond these, the large number of parameters that must be reflected in meaningful formulas will preclude their derivations, simply due to algebraic difficulties.

The numerical analysis outlined in our paper enables an intuitive understanding of a large class of composite plate to be easily gained. The relatively low computational effort must be emphasized. Also, a great deal of information is revealed in a straightforward manner, e.g., the fundamental decay rate as well as the higher order ones and their displacement and stress distributions over the plate's cross section. Although not as elegant as a formula, the numerical technique is equally effective in sorting out the governing parameters on decay rates.

It would be a nice outcome if the numerical studies could lead to relatively simple formulas containing the essential nature of the rate of exponential decay.

A Simple and Effective Pipe Elbow Element – Interaction Effects[1]

J. F. Whatham.[2] I would like to point out that while the experimental points in Figs. 9–13 were taken from Whatham [1] and suitably acknowledged, no mention was made of the analytic solution in that paper. The results from that solution were closer to the experimental points than were the curves presented, particularly for circumferential stress; references [2–5] describe the method more fully.

The Bathe-Almeida model with only two boundary conditions cannot adequately describe the true situation in a curved pipe at a rigid flange where four strain components must vanish.

A computer program package BENDPAC written in FORTRAN IV and ASSEMBLER for an IBM 3031 computer is available from the Australian Atomic Energy Commission,

[1] By S. B. Dong and D. B. Goetschel, and published in the March, 1982, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 129–135.

[2] Professor of Mechanics, Department of Metallurgy, Mechanics and Materials Science, Michigan State University, East Lansing, Mich. 48824.

[1] By K. J. Bathe and C. A. Almeida, and published in the March, 1982 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 165–171.

[2] Australian Atomic Energy Commission, Lucas Heights Research Laboratories, New Ittawarra Road, Lucas Heights, NSW Australia.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and the programs contained therein are designed to give an analytic solution for the following:

(*a*) Stresses and deflections in pipe bends terminated by flanges, infinitely long tangent pipes or short equal length flange-ended tangents when under pressure (PRESEF), or pure in-plane bending (BENDEF).

(*b*) Stresses in flange-ended pipe bends from in-plane end loading other than pure bending (SHEFEF), coil spring type out-of-plane loading (COILEF), or any other type of out-of-plane loading (TURNEF).

(*c*) Flexibility matrices for flange-ended pipe bends under any in-plane end loading (FLEXIN) or out-of-plane end loading (FLEXOT).

Individual programs PRESEF and BENDEF are available from the National Energy Software Center, Argonne National Laboratory, while the total BENDPAC package has been recently supplied to the Center.

References

1 Whatham, J. F., "In-Plane Bending of Flanged Pipe Elbows," *Proceedings, Metal Structures Conference*, the Institution of Engineers, Perth, Australia, Nov. 30-Dec. 1, 1978, also: *Transactions, Institution of Engineers, Australia*, Vol. CE21, No. 2, 1979, pp. 80-85.

2 Whatham, J. F., and Thompson, J. J., "The Bending and Pressurizing of Pipe Bends With Flanged Tangents," *Journal of Nuclear Engineering and Design*, Vol. 54, No. 1, 1979, pp. 17-28.

3 Whatham, J. F., "Thin Shell Equations for Circular Pipe Bends," *Journal of Nuclear Engineering and Design*, Vol. 65, No. 1, 1981, pp. 77-89.

4 Whatham, J. F., "Thin Shell Analysis of Circular Pipe Bends," *Transactions, Institution of Engineers, Australia*, Vol. CE23, No. 4, 1981, pp. 234-245.

5 Whatham, J. F., "Thin Shell Analysis of Non-Circular Pipe Bends," *Journal of Nuclear Engineering and Design*, Vol. 67, No. 2, 1981, pp. 287-296.

Authors' Closure

As was stated, the objective in our research is to obtain a *simple* pipe element that can be used in practical and general analyses of large piping systems – static, dynamic, materially nonlinear, and large displacement analyses of assemblages of many bends and straight sections [1-4]. Clearly, in this development we must balance the simplicity of formulation and solution cost versus the predictive capability of the element, and our primary objective is to predict the flexibility of the structure accurately.

There is no doubt that for single pipe bends with or without straight sections or flanges, more accurate solutions on detailed stress distributions, than obtained with our simple pipe model, can be calculated. For example, finite element idealizations using isoparametric shell elements can be used to obtain very high accuracy solutions for linear and nonlinear conditions [5], or for certain solutions Whatham's programs can be employed.

References

1 Bathe, K. J., and Almeida, C. A., "A Simple and Effective Pipe Elbow Element — Linear Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, 1980, pp. 93-100.

2 Bathe, K. J., and Almeida, C. A., "A Simple and Effective Pipe Elbow Element — Interaction Effects," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 165-171.

3 Bathe, K. J., and Almeida, C. A., "A Simple and Effective Pipe Elbow Element — Pressure Stiffening Effects," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 1982, pp. 914-916.

4 Bathe, K. J., Almeida, C. A., and Ho, L. W., "A Simple and Effective Pipe Elbow Element — Nonlinear Analysis," *J. Computers and Structures*, in press.

5 Bathe, K. J., *Finite Element Procedures in Engineering Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1982.

Characteristic Forms of Differential Equations for Wave Propagation in Nonlinear Media[1]

M. Ziv[2]. I was most surprised to see this paper whose objectives and outline are systematically the same as those in my paper [1]. These two papers seem too similar to be a coincidence. My surprise stems from the fact that my manuscript was first submitted on June 17, 1980 for publication to Professor Ting, acting as an Associate Editor of the JOURNAL OF APPLIED MECHANICS. Based on his own opinion that my analysis is wrong, Ting rejected my manuscript from being published in the JOURNAL OF APPLIED MECHANICS. Yet, a few months later, Ting submitted his paper, without referring to my work, presenting results which are obtainable from my manuscript.

In his paper, Professor Ting claims that his results are equations (21) and (64). But it is immediately seen that his equation (21) is a special case of my result, equation (35) of [1]. Upon a straightforward reduction, my result, equation (17) of [1], also agrees with his equation (64). Obviously, while rejecting my work, Professor Ting has adapted my objectives only to arrive at results already available to him from my manuscript. Yet, no mention of my manuscript is found in his paper.

In order to show that equation (35) of [1] becomes Ting's equation (21) upon reduction, the following changes in representations are made in equation (35) of [1]: the body force is deleted, the distinction between upper case letters and lower case letters is neglected, the symbols become those of Ting's, and finally rectangular Cartesian coordinates are used taking, for example, the x-direction. With these simplifications, equation (25) of [1] now reads

$$\frac{d\sigma_{xx}}{dt} - \rho_0 c \frac{dv_x}{dt} = M_{xxqp} v_{q,p} - \rho_0 c^2 v_{x,x} - c\sigma_{xp,p} + c\sigma_{xx,x} \quad \text{along} \quad dx/dt = c$$

where M_{kqdl} is the term $\dfrac{\partial^2 w}{\partial E_K^N \partial E_D^Q} x^d_{,D} x^N_{l,} + T^K_Q \delta^d_l$

in equation (35) of [1].

When Ting's equation (21) is now written in terms of his equation (15), i.e., $b_i = cn_i$ and the x-direction, for example, is taken, Ting's equation (21) becomes the one written in the foregoing. It was thus shown that equation (21) of Ting's paper is a special case obtained directly from equation (35) of [1]. In the same manner, it can be easily shown that equation (64) of Ting's paper is a special case of equation (17) of [1].

Also, it should be noted that equation (43) of Ting's paper is equation (38) of [2] when taking the direction cosines to be $n_r = 1$ and $n_\theta = 0$. Reference [2] was listed in my manuscript. No reference was made, however, to these works and no new results are given in Ting's paper.

References

1 Ziv, M., "Speeds and Differential Equations of Large Deformation Waves: Hyperelastodynamics and Hypoelastodynamics," *The Journal of the Acoustical Society of America*, Vol. 70, No. 1, July 1981, pp. 218-227.

2 Ziv, M., "Two-Spatial Dimensional Elastic Wave Propagation by the Theory of Characteristics," *Int. J. Solids and Structures*, Vol. 5, 1969, pp. 1135-1151.

[1] By T. C. T. Ting and published in the December, 1981, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 743-748.

[2] Mechanics and Structures Department, School of Engineering and Applied Science, University of California, Los Angeles, Calif., 90024.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Author's Closure

Professor Ziv said that his manuscript was first submitted on June 17, 1980, for publication to the JOURNAL OF APPLIED MECHANICS. As a matter of fact, the same manuscript was submitted one year earlier in June of 1979 to the JOURNAL OF APPLIED MECHANICS and was handled by another Associate Editor. That earlier manuscript was rejected. With little change he submitted the same manuscript on June 17, 1980, and I was assigned to handle the paper.

The paper was rejected because of the following two basic theoretical errors: (*i*) the direction of the bicharacteristic velocity b_i is assumed to be in the direction of the normal to the wave surface n_i, which is not true in general, and (*ii*) the wave speeds c are not determined from the eigenvalues of the acoustic tensor and hence not all equations of motion are satisfied. Point (*i*) was pointed out by me while point (*ii*) was pointed out by my reviewers as well as the reviewers of the earlier manuscript. Professor Ziv simply refused to acknowledge the errors and submitted his manuscript to the *Journal of the Acoustical Society of America*. Notice that he still writes $b_i = c\, n_i$ in his discussion which implies that b_i is in the direction of n_i.

The author did not refer to Professor Ziv's manuscript at the time of writing the paper because (*a*) Professor Ziv's manuscript has not appeared in print and (*b*) the analysis in the manuscript was theoretically incorrect. No purpose would be served by referring to the manuscript and stating that the unpublished manuscript was in error. Fortunately, the manuscript has now appeared in the *Journal of the Acoustical Society of America*. The public can see that the analyses in Ziv's paper and mine are quite different and can judge for themselves whether the analysis in Ziv's paper is sound, let alone his claim that his results are more general than mine.

Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Journal of Applied Mechanics

Book Reviews

Structural Modelling and Optimization. By D. G. Carmichael. Ellis Horwood Ltd., Chichester, England, 1981. 306 Pages. Price $69.95

REVIEWED BY J. E. TAYLOR[1]

This unique book on optimal structural design features a broad, nicely integrated, and well-documented exposition on mathematical methods in optimization. Special attention given to details of organization have resulted in a form and style of presentation that is both unusually orderly and functionally effective. The author states that his book is "aimed at the graduate student and reference text level." It should in fact prove to be unusually useful as a source book for serious analysts in any of the various disciplines where methods of optimization or control theory are applied.

The author's object in preceding the material on optimization with a treatment of modeling is to establish right at the start a well-organized scheme for the representation of structural analysis. Thus the material of Part A of the book serves to fix the technique and nomenclature for analysis in a style identified as an "engineering systems approach." Topics by chapter title are: A Basis for Modelling (multilevel and single level systems representations are discussed), Single Level State Space Modelling, Model Decomposition, and Staged Systems.

A discussion of "design," stated in general terms, is provided in the 25-page second section of the book. Ideas related to design "in a systems sense," and to the "design process" are presented. A definitive statement on the theory of optimal control is given along with a detailed description of the "The Optimal Control Problem." This material reflects the author's uniform commitment to relate his presentations wherever possible to the style and methods of control theory. A basis for categorization of problems is introduced here as well. The categories are identified with the respective form of equations for analysis: ordinary differential, partial differential, algebraic, and difference.

The major substance of the book is comprised of the material in Part C, entitled, "Deterministic Optimization." A rather comprehensive treatment of mathematical methods is provided in chapters titled: Mathematical Programming, Dynamic Programming, Pontryagin's Principle: Variational Calculus, Singular Control, Multicriteria Optimization, Multilevel Optimization, Lyapunov Theory in Design (to accommodate treatment of stability problems), and Energy and "Optimality Criteria" Based Design. Example structural optimization problems are included in most of these chapters.

Treatments in the text are tied together chapter-by-chapter to developments in the literature via the brief discussions furnished in separate sections entitled, Notes, Comments, and Bibliography. Those sections collectively comprise a valuable annotated listing of roughly 800 references.

[1]Professor of Aerospace Engineering, The University of Michigan, Ann Arbor, Mich. 48109.

Recent Developments in Thermomechanics of Solids. Edited by G. Lebon and P. Perzyna. Springer-Verlag, New York, 1981. 415 Pages. Price $43.70

REVIEWED BY W. O. WILLIAMS[2]

This book presents the texts of three courses presented at the International Center for the Mechanical Sciences (CISM) at Udine. The utility of such courses may be questionable; because the audience usually includes people with no knowledge of the subject and also specialists in the subject, the lecturer feels obliged to begin with very elementary ideas and proceed to very specialized ones, satisfying neither audience. With this in mind one can say that the virtue of such lectures is that they may give a flavor of the subject and a view of its breadth and may also serve as an introduction to the specialized literature. The three presentations here, to varying degrees, possess both of these virtues.

K. Wilmanski in the article "Thermodynamic Foundations of Thermoelasticity" gives an (unfortunately) brief sketch of his ideas for thermodynamics of continua based on accessibility and inherited states. He proceeds to describe the axiomatic derivation of the field equations due to Gurtin and Williams and then introduces constitutive relations, using the scheme of Müller to deduce the equations for a rigid heat conductor including dependence on temperature rate. He then discusses the conductivity types of such materials, following Chadwick and Curie, briefly relating these to the type of heat conduction equation (hyperbolic under some natural restrictions). A final section discusses waves in such materials.

P. Persyna in "Thermodynamics of Dissipative Materials" aims toward viscoplastic materials. He begins with another brief exposition of the axioms of the field theory, then goes on to discuss a class of simple materials (within the format of Noll's "New Theory") whose local states are described by configuration and the vaguely described "method of preparation." He then introduces materials whose state is described by internal state variables, dividing these into those related to internal friction and those related to thermoactivated mechanisms. The next section relates these to rate-type theories, which leads to a discussion of the internal variable description of viscoplasticity. The corresponding evolution equations are related to the behavior of dislocation and there is some discussion of experimental results. This section is quite interesting and has many references to the literature. Finally he presents a brief discussion of rate-independent plasticity theory.

G. Leblon in "Variational Principles in Thermomechanics" begins with a not completely precise but very clear description of basic variational calculus. He emphasizes the approximation of variational and quasi-variational

[2]Professor, Department of Mathematics, Carnegie-Mellon University, Schenley Park, Pittsburgh, Pa. 15213.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

problems by various methods. An appendix discusses finite element methods. The chapter on applications in classical mechanics and elasticity is almost too comprehensive. He discusses least action, Hamilton's principle, as well as all of the standard variational principles of infinitesimal elastostatics, elastodynamics, and finite elastostatics. Several examples are given. The chapter on heat conduction is likewise very complete. He discusses the stationary equations, both linear and nonlinear, the nonstationary linear equations, both the standard parabolic equations and the hyperbolic equations, and finally several results for nonstationary nonlinear theory. The chapter on coupled thermoelasticity skims over several variational principles from the recent literature. This set of notes could easily form a text for a graduate engineering course, as the presentation is very concise and thorough.

Instabilities and Catastrophes in Science and Engineering. By J. M. T. Thompson. Wiley, New York, 1982. pp. xvi–226. Price $34.95.

REVIEWED BY H. H. E. LEIPHOLZ[3]

The author of this book is a well-known expert in the field of stability of mechanical systems. Therefore, in writing this work, he was able to draw a wealth of facts and examples from his previous research. Yet, having pursued the development of stability theory in general, i.e., in mathematics and science, and having studied most recent results, for example, in the context of catastrophy theory, he was able to add important results coming from other branches of science to his results in engineering, so that the book has become a broad, most impressive, and stimulating survey of modern stability theory. Should anybody, student or even experienced specialist, like to have a fascinating introduction into the state of the art, the ramifications of stability theory, a hint on the new and promising onset of further development, and a survey on present applications of stability theory to various problems in a set of most diverse fields of engineering and science, he should read this book.

In spite of its broad scope, the book is written sufficiently rigorously so as to satisfy the specialist, and at the same time, sufficiently challenging so as to attract any reader to areas described in the book that should be new to him.

Some of the important modern topics dealt with in the book are: catastrophe theory, stability of nonconservative systems, dynamics of a strange attractor, etc. Fields of applications touched on are: structural engineering, astrophysics, nuclear physics, biochemistry, ecology, hydrodynamics, space mechanics, and neurology. The book is highly recommended to anybody interested in instability phenomena and stability theory.

Mechanics of Brittle Fracture. By G. P. Cherepanov. (Translation by A. L. Peabody, edited by R. DeWit and W. C. Cooley, from the 1974 Russian edition by Nauka Press, with 1977 supplementary material from the author.) McGraw-Hill, New York, 1979. pp. viii–939. Price $97.00.

REVIEWED BY J. R. RICE[4]

This is a translation of the 1974 Russian edition, with supplements provided by the author in 1977. Cherepanov is among the most original and imaginative contributors to the flowering of fracture mechanics theory in the late 1960s and 1970s. This book fairly reflects his remarkably diverse interests and insights.

Crack mechanics, mostly elastic, forms the central theme. There are strong sections on singular fields, plane elasticity solutions to crack problems, energy release-rate calculations, and self-similar elastodynamic solutions, and there is also a brief catalog of elastic crack solutions. In addition there is much contact, if not always developed very completely, with adjacent areas of physics and chemistry, ranging from principles of bonding to optically induced fracturing to ion transport in electrolytic solutions within crack spaces. Supplementary discussions and applications also have a wide scope. For example, there is a discussion of path invariant integrals not only as energy release-rate representations in fracture growth, but also as applied to dielectric phenomena and fluid dynamics. The fracture phenomena and mechanisms discussed include fatigue, environmentally influenced fracture, micromechanisms in fiber-strengthened composites, erosion, drilling, rockburst mechanisms, and much more. There is no shortage of speculation, and one cannot help but wish that the author had been a little more self-critical before jumping to facile explanations of plainly more subtle phenomena, for example, in the fatigue and corrosion cracking areas.

The flavor throughout the book is intensely personal, and one will be disappointed if it is approached as a compendium of all that is useful and permanent in the worldwide development of fracture mechanics. In the foreword to this English translation, Cherepanov warns his readers, and perhaps hopes thereby to deflect some critics, that his works have been ". . . criticized for paying too little attention to the points of view and works of others." He continues: "Unfortunately, I was unable to overcome this shortcoming in the present edition as well, as it represents primarily the results of my own work." The latter quote is certainly accurate, except perhaps that in many places the work could, with more accuracy, be called his synthesis in recapitulation and extension of the results of others. The referencing to others is often offhand, although in aggregate there are 475 references cited (of which 85 are to Cherepanov's papers). The selectivity leaves the book less strong than it could be on elastic-plastic frature phenomena, on unsteady crack dynamics, and on connections with fracture phenomena from the materials science viewpoint, e.g., as developed in the books by Knott and by Lawn and Wilshaw. Similarly, it is less successful than the book by Brock on structural applications. Yet this book is distinctly different in flavor from any of those three, and certainly stronger in basic mechanics.

Elementary Finite Element Method. By C. S. Desai. Prentice-Hall, Englewood Cliffs, N.J. pp. xiv–434. Price $22.95.

REVIEWED BY S. KELSEY[5]

The professed aim of this book is to provide a broad, but practical introduction to the finite element method, suitable for the undergraduate engineering student or other finite element novice. The method is presented, not as a development of matrix methods of structural analysis (a historical view) nor as a technique for the approximate solution of differential equations (a mathematician's view), but as a general and coherent approach to the analysis of physical behavior by discrete modeling.

---

[3] Professor and Chairman, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

[4] Division of Applied Sciences, Harvard University, Cambridge, Mass 02138.

[5] Professor, Department of Civil Engineering, University of Notre Dame, Notre Dame, Ind. 46556.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The essential scheme of the book is to develop understanding of the method and familiarity with its techniques by analyzing successively more complex and sophisticated problems. At the same time, the generality is emphasized by the range of different problems considered. Mathematical formality and rigor are kept to a minimum; in particular the calculus of variations is avoided, in keeping with the assumed mathematical background of the reader. Attention is directed almost exclusively to the generation and solution of the governing equations of the discrete systems. The reader is presumed to possess at least a basic understanding of the physical subjects treated.

While the author rightly emphasizes the value of hand calculation with small examples as a way of acquiring insight, he also emphasizes that it is only through the computer that the finite element method finds its real purpose and application. He is therefore at pains to provide in the book a detailed and well-documented computer program for one-dimensional problems. Information on other programs, applicable to many of the other topics covered in the book and available from the author, is also given.

The book really begins in Chapter 2, after a brief and rather mysterious philosophical disquisition on the idea of discretization, with an introduction of the basic concepts and a skeleton of theory. Not surprisingly, the first specific problem to be treated, in Chapter 3, is the axially loaded, elastic bar. The classic displacement analysis is derived both from the principle of minimum potential energy and from Galerkin's method. In addition, a mixed method, based on the Hellinger-Reissner principle, is presented. One-dimensional steady flow problems are next discussed and the close analogy with the elastic bar emphasized in Chapter 4. The flow problem also provides a vehicle for enlargement of the analysis into time-varying behavior in Chapter 5.

Chapter 6 contains the one-dimensional computer program, applicable to the foregoing models, and in Chapter 7 the bending behavior of beams serves as an example to introduce higher-order continuity conditions (slope as well as displacement) and, as a refinement, interpolation functions with continuity of curvature at nodes. The remaining one-dimensional problems treated are: mass transport (convection/diffusion equation), overland flow, and elastic stress-wave propagation in Chapters 8–10.

Two-dimensional problems are introduced in Chapter 11 via the St. Venant torsion problem. Four analyses are given. Besides the solutions in terms of warping and stress functions, there is a hybrid and a mixed method, all using simple triangular elements. Other two-dimensional field problems (potential flow, heat flow, and seepage) are treated in Chapter 12 to emphasize unity and generality, but the analysis uses simple (quadrilateral) isoparametric elements. Chapter 13 covers plane elasticity, using the displacement method.

In the final chapter (14), the scope of the finite element method and its applicability to heterogeneous systems is illustrated by a discussion of the modeling necessary for a building frame with slab floors and elastic foundations.

The range of technical subjects is very impressive and the versatility of the finite element method is clearly demonstrated. A student working through the book, with its constant reinforcement of ideas and tactics and its widening horizons, and making use of the computer programs available, must very quickly acquire familiarity and competence in basic finite element technology.

But there are some flaws, and these are mainly in basic, theoretical aspects. Thus, the idea of isoparametric elements is presented rather casually, and the "natural" or "intrinsic" coordinates, fundamental to the idea, are merely and misleadingly described as "local" coordinates, with no explanation. It is also difficult to see what the complementary energy method has to offer in the analysis of the axially loaded bar.

More serious, however, is the fundamental question of the variational basis of the finite element method. The author uses a mixture of approaches; sometimes a variational principle is invoked, sometimes the Galerkin method is applied, once the principle of virtual work. Discussion of these and their relationships can only be loose and qualitative without the help of variational calculus. Although the principle of minimum total potential energy, as a condition for equilibrium, should be familiar to an undergraduate from his or her physics training, the other principles used, like the Hellinger-Reissner, must be quite mysterious. And even if these are acceptable for defining element behavior, the question of natural boundary conditions is extremely obscure without the variational argument. A brief concise exposition of the essential variational approach would have added greatly to the strength of the book, by providing a clearer theoretical base to complement the development.

Continuum Mechanics. By A. J. M. Spencer. Longman Mathematical Texts, 1980. 183 Pages. Price $13.50.

REVIEWED BY R. L. FOSDICK[6]

The author has intended this book to be an introduction to the theory of continuum mechanics in a form that is suitable for undergraduate students. I think that his aim has been achieved and that the approach and scope of material covered is reasonable for the engineering undergraduate student at the junior or senior level. In 183 pages and 11 chapters it is not possible to stray too far from the beaten track and by design the author does not do so. The usual elements of the kinematics of motion and deformation, stress, and balance laws are to be found in a general and understandable setting which is not compromised by ad hoc linearity assumptions. There is one chapter that is devoted to linear constitutive relations for elasticity, classical viscous fluids, and viscoelasticity. In addition, there is one chapter that gives a brief, but instructive description of the constitutive theory for nonlinear hyperelasticity, and of the invariance restrictions that lead to the Reiner-Rivlin fluid. Simple materials with general history dependence are merely mentioned, and the basic elements of classical plasticity theory are discussed and formulated in four pages. There are two supporting chapters at the beginning of this book on matrix algebra, vectors, and cartesian tensors and one at the end of the book on cylindrical and spherical coordinates. The notion used throughout the book is a blend of direct or matrix (i.e., index-free) and the indicial for rectangular cartesian coordinates. There is no thermomechanics covered in this book. While there is a reasonable number of exercises on the mechanics topics with an appendix of answers, there are none that deal explicitly with vectors and tensors.

[6]Professor, Department of Aerospace Engineering, University of Minnesota, Minneapolis, Minn. 55455.

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ERRATA

Errata on "On the Logarithmic Singularity of Free-Edge Stress in Laminated Composites Under Uniform Extension," by R. I. Zwiers, T. C. T. Ting, and R. L. Spilker, and published in the September, 1982 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 561–569.

On page 561, the last sentence of the second paragraph in the second column should read: If k^* so obtained happens to be zero at a particular free-edge, there is no r^δ singularity at that free-edge. The purpose of this paper is to provide the exact nature of the singularity at the free-edge.

On page 567, equation (77*a*) should read:

$$u_i^{(p)} = \Sigma\{ (da_L/d\delta)^{(p)} Re(v_{i,L} Z_L) + (d\tilde{a}_L/d\delta)^{(p)} Im(v_{i,L} Z_L) \} + \delta_{i3} x_3 \qquad (77a)$$

On page 568, 7th line from the bottom of the left column should read: $\tilde{\sigma}_{21}{}' = \tilde{\sigma}_{21}$ and $\tilde{\sigma}_{23}{}' = \tilde{\sigma}_{23}$ at $\phi = 0$, only $\tilde{\sigma}_{22}(0)$, $\tilde{\sigma}_{21}(0)$, $\tilde{\sigma}_{23}(0)$

Copyright © 1982 by ASME
Downloaded 02 May 2010 to 171.66.16.246. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm