Transactions of the ASME®

Technical Editor, **LEWIS T. WHEELER (1997)**

Department of Mechanical Engineering,
University of Houston,
Houston, TX 77204-4792

APPLIED MECHANICS DIVISION

Chairman, **C. T. HERAKOVICH**
Secretary, **T. HUGHES**
Associate Technical Editors,
R. ABEYARATNE (1997)
T. R. AKYLAS (1997)
J. L. BASSANI (1999)
R. BECKER (1998)
S. A. BERGER (1997)
I. M. DANIEL (1999)
W. J. DRUGAN (1997)
J. T. JENKINS (1999)
J. W. JU (1998)
V. KINRA (1999)
S. KYRIAKIDES (1997)
S. LICHTER (1998)
W. K. LIU (1999)
X. MARKENSCOFF (1997)
M. ORTIZ (1998)
N. C. PERKINS (1999)
J. N. REDDY (1998)
S. W. SHAW (1997)
M. SHINOZUKA (1997)
M. TAYA (1999)

BOARD ON COMMUNICATIONS

Chairman and Vice-President
R. MATES

OFFICERS OF THE ASME
President, **R. J. GOLDSTEIN**
Executive Director, **D. L. BELDEN**
Treasurer, **J. A. MASON**

PUBLISHING STAFF
Managing Director, Engineering
CHARLES W. BEARDSLEY
Director, Technical Publishing
PHILIP DI VIETRO
Managing Editor, Technical Publishing
CYNTHIA B. CLARK
Managing Editor, Transactions
CORNELIA MONAHAN
Production Coordinator
JUDITH SIERANT
Production Assistant
MARISOL ANDINO

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $210.00 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017.

Periodicals postage paid at New York, NY and additional mailing office. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.

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, $40.00, annually: to nonmembers, $210.00. Add $30.00 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, Para. 3), **COPYRIGHT** © **1996** by The American Society of Mechanical Engineers. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by ASME to libraries and other users registered with the Copyright Clearance Center (CCC). Transactional Reporting Service provided that the base fee of $3.00 per article is paid directly to CCC, Inc., 222 Rosewood Drive, Danvers, MA 01923. Request for special permission or bulk copying should be addressed to Reprints/Permission Department.

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc.
Canadian Goods & Services Tax Registration #126148048.

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 63 • NUMBER 3 • SEPTEMBER 1996

TECHNICAL PAPERS

565 **Bifurcation and Chaos in Externally Excited Circular Cylindrical Shells**
Char-Ming Chin and A. H. Nayfeh

575 **Transient Response in a One-Dimensional Model of Thermoelastic Contact**
Z. S. Olesiak and Yu. A. Pyryev

582 **On Nonuniqueness and Stability in Barber's Model of Thermoelastic Contact**
Z. S. Olesiak and Yu. A Pyryev

587 **A Higher-Order Theory for Vibration of Doubly Curved Shallow Shells**
K. M. Liew and C. W. Lim

594 **Freezing Range Effect on Shell Growth Instability During Alloy Solidification** (96-WA/APM-2)
L. G. Hector, Jr., Woo-Seung Kim, and O. Richmond

603 **Stability of Generalized Equilibria in Constrained Pseudodissipative Systems**
J. A. Walker

610 **Rotation of Clamped Parts due to Reciprocating Motion**
E. S. Edelstein and J. J. Blech

621 **The Debonded Interface Anticrack**
X. Markenscoff and Luqun Ni

628 **An Exact Three-Dimensional Solution for Simply Supported Rectangular Piezoelectric Plates**
P. Bisegna and F. Maceri

639 **The Matrix-Fiber Crack in an Elastic Solid**
A. P. S. Selvadurai, B. M. Singh, and M. C. Au

650 **Consistent Infinitesimal Finite Element Cell Method: Three-Dimensional Scalar Wave Equation**
J. P. Wolf and Chongmin Song

655 **The Wrinkling of Rectangular Webs Under Nonlinearly Distributed Edge Loading**
C. C. Lin and C. D. Mote, Jr.

660 **Anisotropic Beam Theories With Shear Deformation**
H. Murakami, E. Reissner, and J. Yamakawa

669 **Electrorheological Dampers, Part I: Analysis and Design** (96-WA/APM-5)
H. P. Gavin, R. D. Hanson, and F. E. Filisko

676 **Electrorheological Dampers, Part II: Testing and Modeling** (96-WA/APM-6)
H. P. Gavin, R. D. Hanson, and F. E. Filisko

683 **Rotation of a Clamped Spherical Ball due to Linear Reciprocating Motion**
E. S. Edelstein and J. J. Blech

692 **Application of Objective Rates in Mechanical Modeling of Solids**
W. D. Reinhardt and R. N. Dubey

699 **Distortion and Residual Stress in Quenched Aluminum Bars**
R. Becker, M. E. Karabin, J. C. Liu, and R. E. Smelser

706 **Micromechanical Definition of the Strain Tensor for Granular Materials**
N. P. Kruyt and L. Rothenburg

712 **Vertical Vibration of a Deep Bed of Granular Material in a Container**
C. R. Wassgren, C. E. Brennen, and M. L. Hunt

720 **Modeling of Cyclic Ratchetting Plasticity, Part I: Development of Constitutive Relations**
Y. Jiang and H. Sehitoglu

726 **Modeling of Cyclic Ratchetting Plasticity, Part II: Comparison of Model Simulations With Experiments**
Y. Jiang and H. Sehitoglu

734 **Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part I: Elastic Response**
D. L. Littlefield

742 **Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part II: Plastic Response**
D. L. Littlefield

750 **Scaled Memory Description of Hysteretic Material Behavior**
J. P. Bardet

758 **Simulation of Multivariate Gaussian Fields Conditioned by Realizations of the Fields and Their Derivatives**
Y. J. Ren, I. Elishakoff, and M. Shinozuka

(Contents continued on Outside Back Cover)

Downloaded 04 May 2010 to 171.66.16.52. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

CONTENTS (CONTINUED)

766 **Stepwise Loading of Half-Spaces in Elliptical Contact** (96-WA/APM-1)
J. Jäger

774 **The Wrinkling of Thin, Flat, Rectangular Webs**
C. C. Lin and C. D. Mote, Jr.

780 **An Approximate Solution for the Compression of a Bonded Thin Annular Disk**
Yun Ling

788 **Intergranular Crack Nucleation in Bicrystalline Materials Under Fatigue**
H. M. Shodja, Y. Hirose, and T. Mura

796 **Diffusive Crack Growth at a Bimaterial Interface** (96-WA/APM-7)
Tze-jer Chuang, June-Liang Chu, and Sanboh Lee

804 **The Sliding Velocity Flow of Rough Collisions in Multibody Systems**
J. A. Batlle

810 **An Energy-Based Approach to Computing Resonant Nonlinear Normal Modes**
M. E. King and A. F. Vakakis

820 **On the Nonuniqueness of Solutions Obtained With Simplified Variational Principles**
H. Mang, P. Helnwein, and R. H. Gallagher

828 **Feedback Control of Karman Vortex Shedding**
M. D. Gunzburger and H. C. Lee

836 **Cavity-Opening Tensor Method and its Application to Reflection and Transmission of Ultrasonic Wave by Defects**
Yonglin Xu

BRIEF NOTES

843 Comment on Hyperbolicity of Wave Problem for Valanis' Global Damage Theory
G. Pijaudier-Cabot and Z. P. Bažant

845 Hypersingular Integral Equation for a Curved Crack Problem of Circular Region in Antiplane Elasticity
Y. Z. Chen and Norio Hasebe

ANNOUNCEMENTS AND SPECIAL NOTES

586 Thermal Stresses '97—Announcement

620 68th Annual Meeting of the Society of Rheology—Announcement

705 Change of Address form

835 3rd EUROMECH Solid Mechanics Conference—Announcement

850 Worldwide Mechanics Meetings Conference Listing

852 Information for Authors

ISSN 0021-8936 JOURNAL CODE JE ISSUE CODE 9603

Downloaded 04 May 2010 to 171.66.16.52. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Bifurcation and Chaos in Externally Excited Circular Cylindrical Shells

Char-Ming Chin
Research Associate,
Mem. ASME.

A. H. Nayfeh
University Distinguished Professor,
Fellow ASME.

Department of Engineering Science
and Mechanics,
Virginia Polytechnic Institute and
State University,
Blacksburg, VA 24061-0219

The nonlinear response of an infinitely long cylindrical shell to a primary excitation of one of its two orthogonal flexural modes is investigated. The method of multiple scales is used to derive four ordinary differential equations describing the amplitudes and phases of the two orthogonal modes by (a) attacking a two-mode discretization of the governing partial differential equations and (b) directly attacking the partial differential equations. The two-mode discretization results in erroneous solutions because it does not account for the effects of the quadratic nonlinearities. The resulting two sets of modulation equations are used to study the equilibrium and dynamic solutions and their stability and hence show the different bifurcations. The response could be a single-mode solution or a two-mode solution. The equilibrium solutions of the two orthogonal third flexural modes undergo a Hopf bifurcation. A combination of a shooting technique and Floquet theory is used to calculate limit cycles and their stability. The numerical results indicate the existence of a sequence of period-doubling bifurcations that culminates in chaos, multiple attractors, explosive bifurcations, and crises.

1 Introduction

Due to the complete circular symmetry of cylindrical shells, each flexural natural frequency corresponds to two orthogonal mode shapes. The mode with the same spatial variation as the excitation is called the driven mode, whereas the other orthogonal mode is called the companion mode. The first study of the interaction between the driven and companion modes in thin structures was conducted by Evensen (1966). He used the method of averaging to study the influence of one-to-one internal resonances on the inextensional forced flexural response of thin circular rings. He found that the response involves either a single bending mode or two coupled bending modes. He reported experimental observations of nonsteady vibrations and verified them by using analog-computer simulations. Chen and Babcock (1975) used the Lindstédt-Poincaré technique to study the nonlinear response of finite cylindrical shells to harmonic excitations. They studied the interaction of driven and companion modes. They reported experimental observations of "nonstationary" responses "in which the amplitude drifts from one value to another". Maewal (1981) studied the interaction between backward and forward traveling waves in rotating rings. Using a nonlinear inextensionality constraint, Maewal (1986) studied the modal interaction between driven and companion modes of axisymmetric shells. He used the method of multiple scales to derive a set of evolution equations for the amplitudes and phases of the modes. He noted that these equations have the same form as those derived earlier by Miles (1984) for internally resonant surface waves in a circular cylinder. He numerically integrated these evolution equations and showed that, for certain ranges of the excitation frequency, the response is chaotically modulated.

Maganty and Bickford (1987) used the method of multiple scales to study the nonlinear free vibrations of circular rings in the presence of a one-to-one internal resonance between an in-plane mode and an out-of-plane mode. They found a continuous exchange of energy between the coupled modes. Later, (Maganty and Bickford, 1988) they studied the forced response of the same system to a harmonic in-plane or out-of-plane excitation. They showed that for certain values of the excitation frequency, an in-plane excitation can produce nonplanar oscillations and vice versa.

Raouf and Nayfeh (1990) analyzed the nonlinear dynamic response of an infinitely long circular cylindrical shell to a primary-resonant excitation of the nth flexural mode. They directly attacked the continuous system by using the method of multiple scales and derived a set of modulation equations that describe the slow-time evolution of the amplitudes and phases of the interacting modes.

In this paper, we first modify the results derived by Raouf and Nayfeh (1990) and show that only two, instead of three, independent nonlinear coefficients exist in the modulation equations, which reflect the inherent symmetry characteristics of the shell. Second, we compare the results derived by directly attacking the governing partial differential equations with those derived from a two-mode discretization. We show that the discretization approach fails to account for the spatial variation at second order and completely misses the influence of the quadratic nonlinearity on the interaction coefficients. Consequently, the two approaches lead to different bifurcations.

2 Direct Approach

Following McIvor (1966) and Raouf and Nayfeh (1990), we write the equations of motion governing the dynamic response of an infinitely long cylindrical shell to a primary-resonant excitation of one of its two orthogonal flexural modes as

$$\ddot{w} + w - \psi' + \alpha^2(w^{iv} + 2w'' + w)$$

$$= -\dot{\psi}^2 - \frac{1}{2}w'^2 - 2w\psi' + \psi'^2 + \psi' w'' - ww'' + w'\psi''$$

$$+ w\dot{\psi}^2 - 2w'\psi''\psi' + \frac{1}{2}w'^2\psi' - w''\psi'^2 + \frac{3}{2}w''w'^2$$

$$- w\psi'^2 + ww'\psi'' - ww'^2 + ww''\psi' - w^2w''$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Apr. 21, 1995; final revision, Oct. 16, 1995. Associate Technical Editor: S. Kyriakides.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+\alpha^2\Big(-\frac{11}{2}w'^2+4w''\psi'+5w''\psi'''+4\psi''w'-6w''^2$$

$$+8w'''\psi''-8w'''w'+4w^{iv}\psi'+w\psi'''-11ww''$$

$$-4ww^{iv}-3w^2+w'\psi^{iv})$$

$$+\frac{a(1-\nu^2)}{Eh}P(1+\psi'-w) \quad (1)$$

$$\ddot{\psi}-\psi''+w'=2w\ddot{\psi}+2\dot{w}\dot{\psi}+w'w''-2w'\psi'-2w\psi''$$

$$+2ww'+ww''w'+w^2\psi''-w'^2\psi''+\tfrac{1}{2}w'^3+2ww'\psi'$$

$$-2w''w'\psi'-2w\dot{w}\dot{\psi}-w^2\ddot{\psi}$$

$$+\alpha^2(w''w'-w''w'''-ww'''+w^{iv}w') \quad (2)$$

where the overdot indicates the partial derivative with respect to time t, the prime indicates the partial derivative with respect to the angle θ, the radial displacement w is made nondimensional by using the undeformed radius a of the shell, ψ is the angular displacement, time is made nondimensional by using the characteristic time $a\sqrt{\rho(1-\nu^2)/E}$, ρ is the density of the shell per unit width, ν is Poisson's ratio, E is Young's modulus, h is the thickness of the shell, and $\alpha^2=h^2/12a^2$. Here one of the two orthogonal nth modes is excited by a primary-resonant pressure P of the form

$$\frac{a(1-\nu^2)}{Eh}P=\frac{1}{2}\epsilon^3 f\cos n\theta e^{i\Omega t}+cc, \quad (3a)$$

$$\Omega=\omega_n+\epsilon^2\sigma \quad (3b)$$

where cc stands for the complex conjugate of the preceding terms, ϵ is a small nondimensional parameter, ω_n is the natural frequency of the nth mode, and σ is a detuning parameter. It follows from Eq. (3a) that the cos $n\theta$ mode is the driven mode and the sin $n\theta$ mode is the companion mode.

Using the method of multiple scales (Nayfeh, 1973, 1981) to attack directly Eqs. (1) and (2), following steps similar to those used by Raouf and Nayfeh (1990), and taking into account the interaction of the excited nth flexural mode with its companion mode, we obtain

$$w=\epsilon[A_1(t)\cos n\theta+A_2(t)\sin n\theta]e^{i\omega_n t}$$

$$+\frac{\epsilon^2}{\Delta_n}[4c_1(n^2-\omega_n^2)+2nd_1][(A_1^2-A_2^2)\cos 2n\theta$$

$$+2A_1A_2\sin 2n\theta]e^{2i\omega_n t}+\frac{\epsilon^2}{\Theta_n}[2n^2c_3+nd_2]$$

$$\times[(A_2\bar{A}_2-A_1\bar{A}_1)\cos 2n\theta-2A_2\bar{A}_1\sin 2n\theta]$$

$$+\frac{\epsilon^2c_2}{1+\alpha^2-4\omega_n^2}(A_1^2+A_2^2)e^{2i\omega_n t}$$

$$-\frac{\epsilon^2c_4}{1+\alpha^2}(A_1\bar{A}_1+A_2\bar{A}_2)+cc+\ldots \quad (4)$$

$$\psi=\epsilon[\Gamma_nA_1(t)\sin n\theta-\Gamma_nA_2(t)\cos n\theta]e^{i\omega_n t}$$

$$+\frac{\epsilon^2}{\Delta_n}[(1-4\omega_n^2+\alpha^2(4n^2-1)^2)d_1+2_{nc_1}]$$

$$\times[(A_1^2-A_2^2)\sin 2n\theta-2A_1A_2\cos 2n\theta]e^{2i\omega_n t}$$

$$+\frac{\epsilon^2}{\Theta_n}\left[\frac{1}{2}(1+\alpha^2(4n^2-1)^2)d_2+nc_3\right][(A_2\bar{A}_2$$

$$-A_1\bar{A}_1)\sin 2n\theta+2A_2\bar{A}_1\cos 2n\theta]+cc+\ldots \quad (5)$$

$$2i\dot{A}_1=8\epsilon^2\Lambda_1A_1(A_1\bar{A}_1+A_2\bar{A}_2)$$

$$+8\epsilon^2\Lambda_2A_2(A_2\bar{A}_1-A_1\bar{A}_2)+\epsilon^2Fe^{i\epsilon^2\sigma t} \quad (6)$$

$$2i\dot{A}_2=8\epsilon^2\Lambda_1A_2(A_1\bar{A}_1+A_2\bar{A}_2)$$

$$-8\epsilon^2\Lambda_2A_1(A_2\bar{A}_1-A_1\bar{A}_2). \quad (7)$$

The ω_n, Γ_n, Δ_n, Θ_n, c_j, d_j, Λ_j, and F are defined in the Appendix. The solution given by Eqs. (4)–(7) is a modified version of that of Raouf and Nayfeh (1990). Equations (4) and (5) are the same as those obtained by Raouf and Nayfeh (1990). However, the modulation Eqs. (6) and (7) exhibit explicitly the symmetry of the system and the interaction coefficients include some terms that are missing from the results of Raouf and Nayfeh (1990).

Had we used the shell thickness to nondimensionalize the displacement, we would have obtained Eqs. (4)–(7) with A_n being replaced with $\hat{A}_n$, where $\hat{A}_n=aA_n/h$.

3 Discretization

To discretize the problem, we find it more convenient to start from the Lagrangian of the system in nondimensional form; that is

$$\mathcal{L}=\frac{Eah}{2(1-\nu^2)}\int_0^{2\pi}[\dot{w}^2+(1-w)^2\dot{\psi}^2]d\theta$$

$$-\frac{Eah}{2(1-\nu^2)}\int_0^{2\pi}[\psi'^2-2\psi'w+w^2+\psi'w'^2$$

$$+2\psi'w^2-2w\psi'^2-ww'^2-\psi'^2w'^2+w\psi'w'^2$$

$$-w^2w'^2+\frac{1}{4}w'^4+w^2\psi'^2]d\theta$$

$$-\frac{Eah\alpha^2}{2(1-\nu^2)}\int_0^{2\pi}[w''^2+2w''w+w^2+w'^2w$$

$$-4\psi'w''w-2ww'\psi''+w'^2w''-4w''^2\psi'$$

$$-2w'w''\psi''+2w^3+6w^2w''+4ww''^2]d\theta$$

$$+\frac{1}{2}a^2\int_0^{2\pi}P[\psi'-2w-2w\psi'+w^2]d\theta. \quad (8)$$

Assuming that the response consists of the nth driven flexural mode and its companion mode, we let

$$w(\theta,t)=\eta(t)\cos n\theta+\zeta(t)\sin n\theta \quad (9)$$

$$\psi(\theta,t)=\Gamma_n\eta(t)\sin n\theta-\Gamma_n\zeta(t)\cos n\theta \quad (10)$$

where Γ_n is defined in Eq. (A2) and $\eta(t)$ and $\zeta(t)$ are the generalized coordinates of the two orthogonal modes. The radial load is taken in the form

$$\frac{a(1-\nu^2)}{Eh}P=\hat{F}(t)\cos n\theta. \quad (11)$$

Comparing Eqs. (9) and (10) with Eqs. (4) and (5), we conclude the assumed discretized solution given by Eqs. (9) and (10) is not correct. Substituting Eqs. (9)–(11) into Eq. (8) and performing the integrations, we obtain

$$\frac{(1-\nu^2)\mathcal{L}}{Eah\pi}=\frac{1}{2}(1+\Gamma_n^2)(\dot{\eta}^2+\dot{\zeta}^2)+\frac{1}{8}\Gamma_n^2(\eta^2\dot{\eta}^2$$

$$+3\zeta^2\dot{\eta}^2-4\eta\zeta\dot{\eta}\dot{\zeta}+\zeta^2\dot{\zeta}^2+3\eta^2\dot{\zeta}^2)$$

$$-\frac{1}{2}(\Gamma_nn-1)^2(\eta^2+\zeta^2)-\frac{1}{8}\delta_n(\eta^2+\zeta^2)^2$$

$$-\frac{1}{2}\alpha^2(n^2-1)^2(\eta^2+\zeta^2)+\left(\frac{1}{2}\Gamma_nn-1\right)\hat{F}(t)\eta \quad (12)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Nonlinear interaction coefficients

nth mode	ω_n	Direct Λ_1	Direct Λ_2	Discretization $\tilde{\Lambda}_1$	Discretization $\tilde{\Lambda}_2$
2	0.03881	−0.02817	0.00291	−42.514	−14.174
3	0.10976	0.10874	0.45535	−84.146	−28.052
4	0.21045	3.6891	5.0364	−150.02	−50.012
5	0.34034	33.138	36.726	−236.52	−78.845

where

$$\delta_n = \Gamma_n n^3 - \Gamma_n^2 n^4 + 3\Gamma_n^2 n^2 + \tfrac{3}{4} n^4 - n^2. \quad (13)$$

Writing down the Euler-Lagrange equations corresponding to the Lagrangian in Eq. (12), we have

$$\ddot{\eta} + \omega_n^2 \eta = -\frac{\delta_n}{2(1+\Gamma_n^2)} \eta(\eta^2 + \zeta^2) - \frac{\Gamma_n^2}{4(1+\Gamma_n^2)} (\eta\dot{\eta}^2 + 6\zeta\dot{\eta}\dot{\zeta} - 5\eta\dot{\zeta}^2 + \eta^2\ddot{\eta} + 3\zeta^2\ddot{\eta} - 2\eta\zeta\ddot{\zeta}) + \frac{\Gamma_n n - 2}{2(1+\Gamma_n^2)} \hat{F}(t) \quad (14)$$

$$\ddot{\zeta} + \omega_n^2 \zeta = -\frac{\delta_n}{2(1+\Gamma_n^2)} \zeta(\eta^2 + \zeta^2) - \frac{\Gamma_n^2}{4(1+\Gamma_n^2)} \times (\zeta\dot{\zeta}^2 + 6\eta\dot{\zeta}\dot{\eta} - 5\zeta\dot{\eta}^2 + \zeta^2\ddot{\zeta} + 3\eta^2\ddot{\zeta} - 2\zeta\eta\ddot{\eta}). \quad (15)$$

Using the method of multiple scales (Nayfeh, 1973, 1981), we find that, to the second approximation, the solution of Eqs. (14) and (15) can be expressed as

$$\eta = \epsilon A_1(t) e^{i\omega_n t} + cc \quad \text{and} \quad \zeta = \epsilon A_2(t) e^{i\omega_n t} + cc \quad (16)$$

where

$$2i\dot{A}_1 = 8\epsilon^2 \tilde{\Lambda}_1 A_1 (A_1\bar{A}_1 + A_2\bar{A}_2) + 8\epsilon^2 \tilde{\Lambda}_2 A_2 (A_2\bar{A}_1 - A_1\bar{A}_2) + \epsilon^2 F e^{i\epsilon^2\sigma t} \quad (17)$$

$$2i\dot{A}_2 = 8\epsilon^2 \tilde{\Lambda}_1 A_2 (A_1\bar{A}_1 + A_2\bar{A}_2) - 8\epsilon^2 \tilde{\Lambda}_2 A_1 (A_2\bar{A}_1 - A_1\bar{A}_2) \quad (18)$$

$$8\omega_n(1+\Gamma_n^2)\tilde{\Lambda}_1 = \tfrac{1}{2}\Gamma_n^2\omega_n^2 + \tfrac{3}{2}\Gamma_n^2 n^4 - \tfrac{9}{2}\Gamma_n^2 n^2 - \tfrac{3}{2}\Gamma_n n^3 - \tfrac{9}{8} n^4 + \tfrac{3}{2} n^2 \quad (19)$$

$$8\omega_n(1+\Gamma_n^2)\tilde{\Lambda}_2 = -\tfrac{5}{2}\Gamma_n^2\omega_n^2 + \tfrac{1}{2}\Gamma_n^2 n^4 - \tfrac{3}{2}\Gamma_n^2 n^2 - \tfrac{1}{2}\Gamma_n n^3 - \tfrac{3}{8} n^4 + \tfrac{1}{2} n^2. \quad (20)$$

Substituting Eqs. (16) into Eqs. (9) and (10), we obtain

$$w = \epsilon[A_1(t)\cos n\theta + A_2(t)\sin n\theta] e^{i\omega_n t} + cc \quad (21)$$

$$\psi = \epsilon[\Gamma_n A_1(t)\sin n\theta - \Gamma_n A_2(t)\cos n\theta] e^{i\omega_n t} + cc. \quad (22)$$

Comparing Eqs. (21) and (22), obtained by attacking the discretized problem, with Eqs. (4) and (5), obtained by attacking directly the governing partial differential equations, we conclude that the former approach leads to erroneous results because it misses the spatial distribution at second order. Moreover, although the modulation Eqs. (17) and (18), obtained by attacking the discretized problem, have the same form as the modulation Eqs. (6) and (7), obtained by directly attacking the governing partial differential equations, the interaction coefficients in both sets of equations are different because the spatial influence of the quadratic terms was missed in the discretized approach. As a result, the equilibrium solutions as well as the dynamic solutions and their stability predicted by the discretization approach are erroneous, as shown in the next section.

The discretized equations fail to describe the dynamics of the partial differential equations because the linear spatial solutions are assumed a priori to describe the spatial solutions of the nonlinear problem. No such assumption is used when attacking the partial differential equations. The components of the nonlinearity in the nonlinear problem that are orthogonal to the linear mode shapes are discarded during the discretization. The magnitude of the error incurred by discretization varies from system to system. In the present case, all of the quadratic

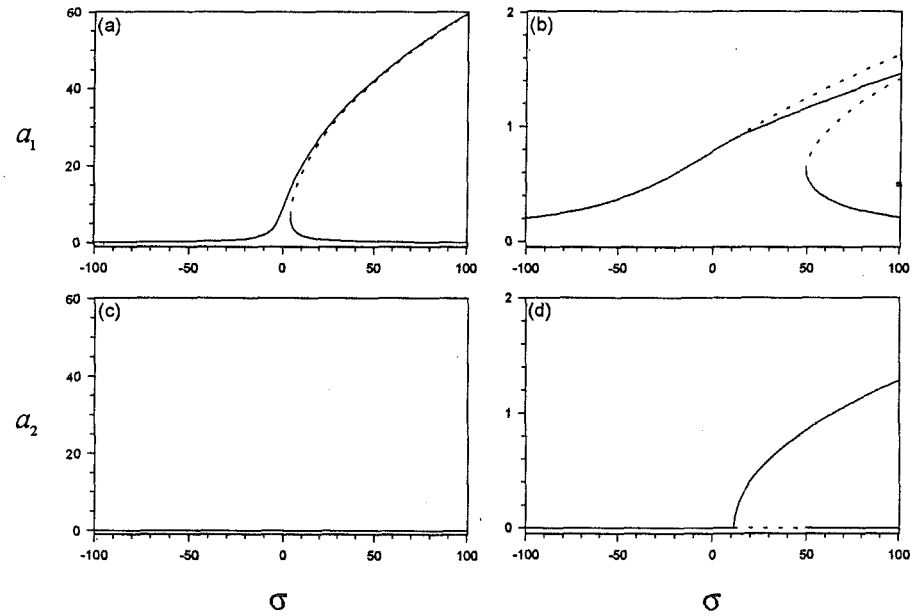


Fig. 1 Frequency-response curves for the second flexural modes with $\omega_2 = 0.03881$, $\mu_n = 0.05$, $\alpha^2 \simeq 2.0918 \times 10^{-4}$, and $f = 2.0$. $a_1 = \sqrt{p_1^2 + q_1^2}$ and $a_2 = \sqrt{p_2^2 + q_2^2}$. Parts (*a*) and (*c*) are obtained from directly attacking the partial differential equations ($\Lambda_1 = -0.02817$, $\Lambda_2 = 0.00291$), whereas (*b*) and (*d*) are obtained from the two-mode discretization approach ($\tilde{\Lambda}_1 = -42.514$, $\tilde{\Lambda}_2 = -14.174$). Solid (dotted) lines denote the stable (unstable) equilibrium solutions.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

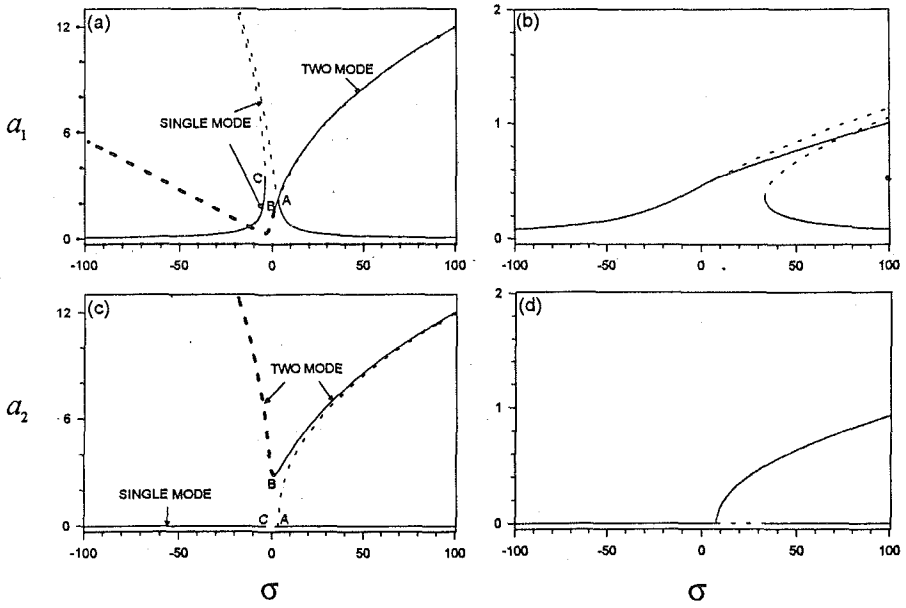


Fig. 2 Frequency-response curves for the third flexural modes with $\omega_3 = 0.10976$, $\mu_n = 0.05$, $\alpha^2 \simeq 2.0918 \times 10^{-4}$, and $f = 2.0$. $a_1 = \sqrt{p_1^2 + q_1^2}$ and $a_2 = \sqrt{p_2^2 + q_2^2}$. Parts (*a*) and (*c*) are obtained from directly attacking the partial differential equations ($\Lambda_1 = 0.10874$, $\Lambda_2 = 0.45535$), whereas (*b*) and (*d*) are obtained from the two-mode discretization approach ($\tilde{\Lambda}_1 = -84.146$, $\tilde{\Lambda}_2 = -28.052$). Solid (dotted) lines denote the stable (unstable) equilibrium solutions and thick-dotted lines denote the unstable foci.

terms in the partial differential equations are discarded during the discretization and hence the error resulting from the discretization depends on the magnitude of the quadratic terms.

4 Stability of Equilibrium Solutions

It follows from the modulation Eqs. (6) and (7) that single as well as two-mode solutions are possible. To determine the stability of these solutions, we introduce the Cartesian transformation

$$A_k = \tfrac{1}{2}\,[p_k(\epsilon^2 t) - iq_k(\epsilon^2 t)]e^{i\epsilon^2\sigma t}, \quad k = 1, 2 \qquad (23)$$

into Eqs. (6) and (7), add the linear modal viscous damping terms $\mu_n p_k$ and $\mu_n q_k$, and obtain

$$p_1' = -\mu_n p_1 - \sigma q_1 - \Lambda_1 q_1(p_1^2 + q_1^2 + p_2^2 + q_2^2) + 2\Lambda_2 p_2(q_1 p_2 - p_1 q_2) \qquad (24)$$

$$q_1' = -\mu_n q_1 + \sigma p_1 + \Lambda_1 p_1(p_1^2 + q_1^2 + p_2^2 + q_2^2) + 2\Lambda_2 q_2(q_1 p_2 - p_1 q_2) + F \qquad (25)$$

$$p_2' = -\mu_n p_2 - \sigma q_2 - \Lambda_1 q_2(p_1^2 + q_1^2 + p_2^2 + q_2^2) - 2\Lambda_2 p_1(q_1 p_2 - p_1 q_2) \qquad (26)$$

$$q_2' = -\mu_n q_2 + \sigma p_2 + \Lambda_1 p_2(p_1^2 + q_1^2 + p_2^2 + q_2^2) - 2\Lambda_2 q_1(q_1 p_2 - p_1 q_2) \qquad (27)$$

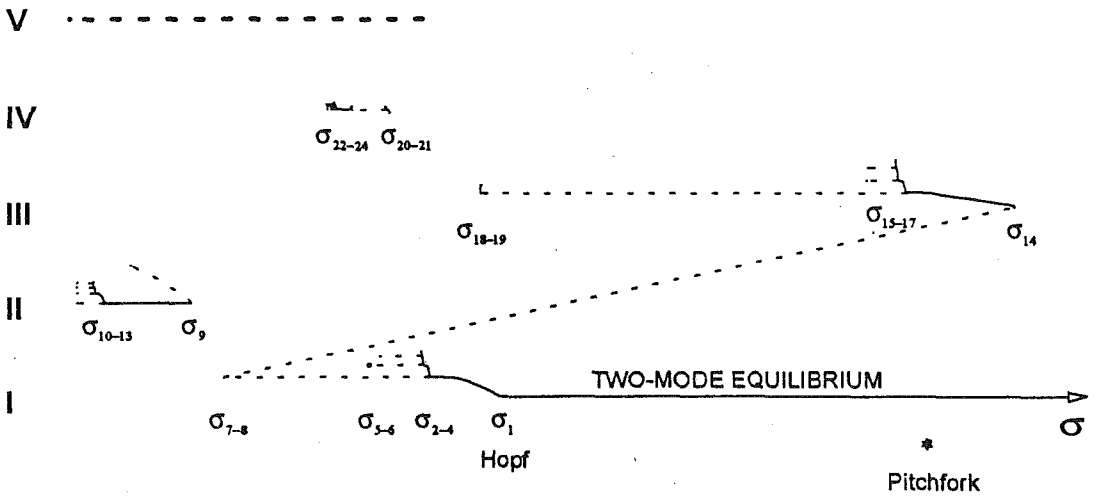


Fig. 3 A schematic bifurcation diagram illustrating the various stages leading to chaos in the neighborhood of the unstable foci, where on branch I: σ_1 (supercritical Hopf) = 1.13953, σ_2 (*p*-2) = 0.7570, σ_3 (*p*-4) = 0.7163, σ_4 (*p*-8) = 0.71148, σ_5 (reverse *p*-8) = 0.48046, σ_6 (reverse *p*-4) = 0.41815, σ_7 (reverse *p*-2) = −0.36348, and σ_8 (cyclic-fold bifurcation) = −0.36786; on branch II: σ_9 (cyclic-fold bifurcation) = −0.54697, σ_{10} (symmetry-breaking) = −1.0081, σ_{11} (*p*-2) = −1.0569, σ_{12} (*p*-4) = −1.0643, and σ_{13} (*p*-8) = −1.0652; on branch III: σ_{14} (cyclic-fold bifurcation) = 3.93572, σ_{15} (*p*-2) = 3.3515, σ_{16} (*p*-4) = 3.3057, σ_{17} (*p*-8) = 3.2999, σ_{18} (reverse *p*-2) = 1.029834, and σ_{19} (jump) = 1.029832; on branch IV: σ_{20} (cyclic-fold bifurcation) = 0.52681, σ_{21} (*p*-2) = 0.50579, σ_{22} (reverse *p*-2) = 0.318, and σ_{23} (*p*-2) = 0.2303, σ_{24} (*p*-4) = 0.22608; and branch V indicates the region of large chaotic attractors.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

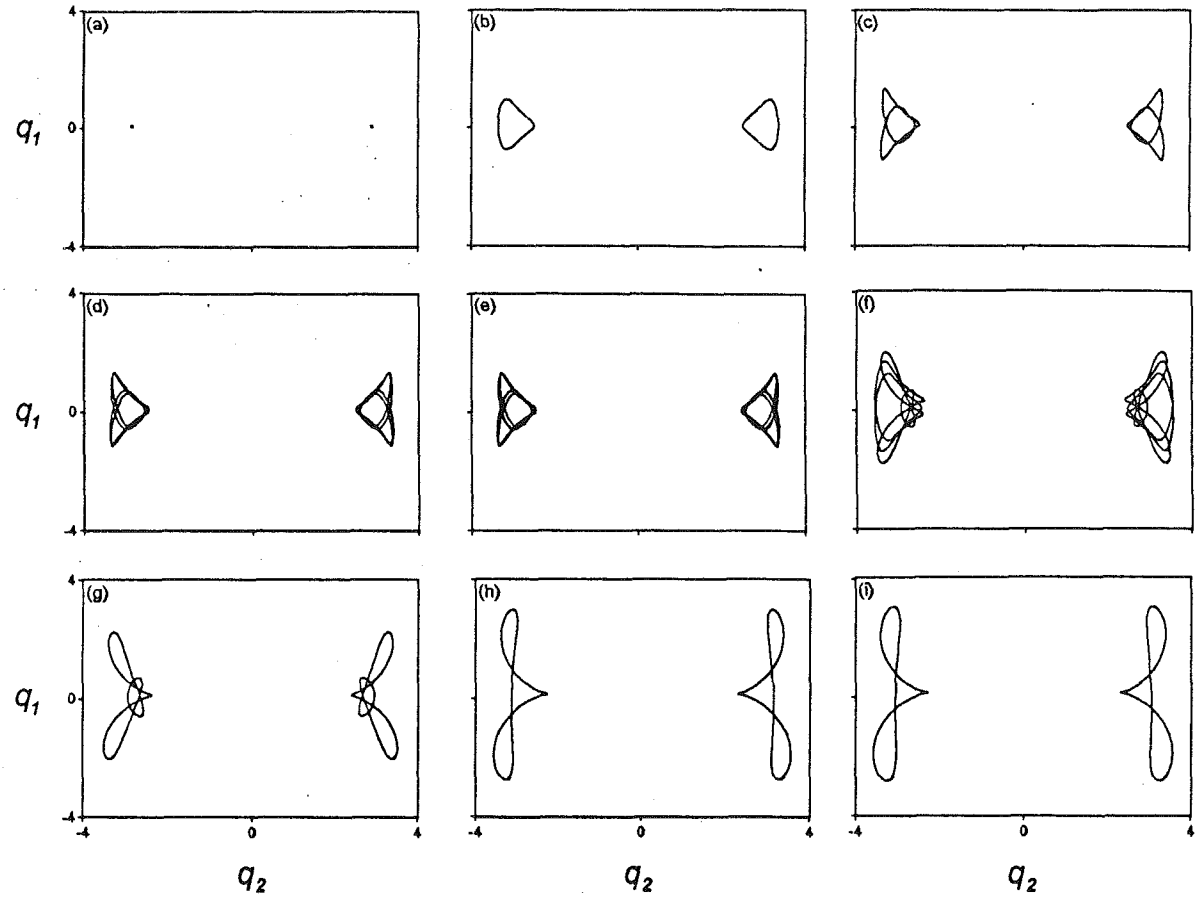


Fig. 4 Two-dimensional projections of the phase portraits onto the q_1–q_2 plane at σ = (*a*) 1.13883, (*b*) 0.75713, (*c*) 0.71633, (*d*) 0.71158, (*e*) 0.71097, (*f*) 0.48046, (*g*) 0.41773, (*h*) −0.36349, and (*i*) −0.36786 on branch I. They show a sequence of period-doubling bifurcations leading to chaos followed by a reverse but fragmented period-doubling sequence of bifurcations.

where the prime indicates the derivative with respect to $\epsilon^2 t$. We note that the modulation Eqs. (24) – (27) are independent of ϵ. Hence, the present results are independent of the actual value of ϵ as long as it is small.

We numerically solve for the equilibrium solutions and find the corresponding eigenvalues for two flexural modes, corresponding to $n = 2$ and $n = 3$, with $\mu_n = 0.05$, $\alpha^2 \simeq 2.0918 \times 10^{-4}$, and $f = 2.0$. It follows from Table 1 that the sign of Λ_1 changes from negative to positive as n increases from $n = 2$ to $n = 3$. This sign change plays an important role in the qualitative behavior of the shell response.

We show the frequency-response curves obtained by using the direct approach and the two-mode discretization approach in Figs. 1 and 2, respectively. For the case of $n = 2$, the response curves predicted by the direct approach, as shown in Figs. 1(*a*) and (*c*), exhibit a hardening-spring type behavior. Moreover, only single-mode solutions are possible, in which only the driven mode responds with an amplitude $a_1 \neq 0$, while the companion mode has a zero amplitude (i.e., $a_2 = 0$). Furthermore, it follows from Figs. 1(*b*) and (*d*) that the discretization approach underestimates the response amplitudes and erroneously predicts the occurrence of a supercritical pitchfork bifurcation (i.e., one of the corresponding eigenvalues crosses the imaginary axis along the real axis from the left to the right-half plane) as σ increases from a small value. As a result, the single-mode solution gives way to a two-mode stable equilibrium solution. In Fig. 2 for the case of $n = 3$, a softening-spring type behavior, as shown in parts (*a*) and (*c*), is obtained by directly

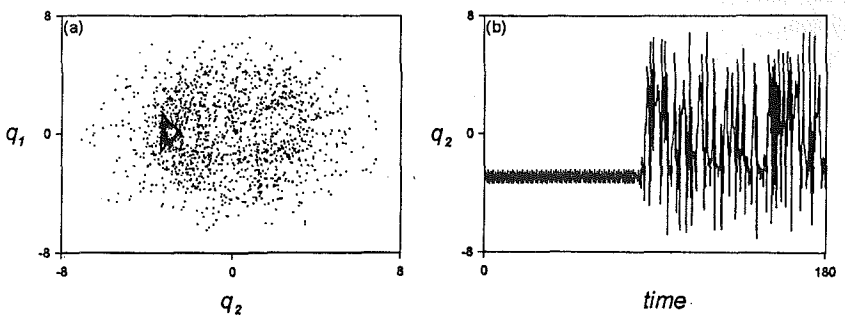


Fig. 5(*a*) A two-dimensional projection of the trajectory onto the q_1–q_2 plane (only shows discrete data points) and (*b*) the time history of q_2 at $\sigma \simeq 0.71061$. They show an explosive type of bifurcation.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

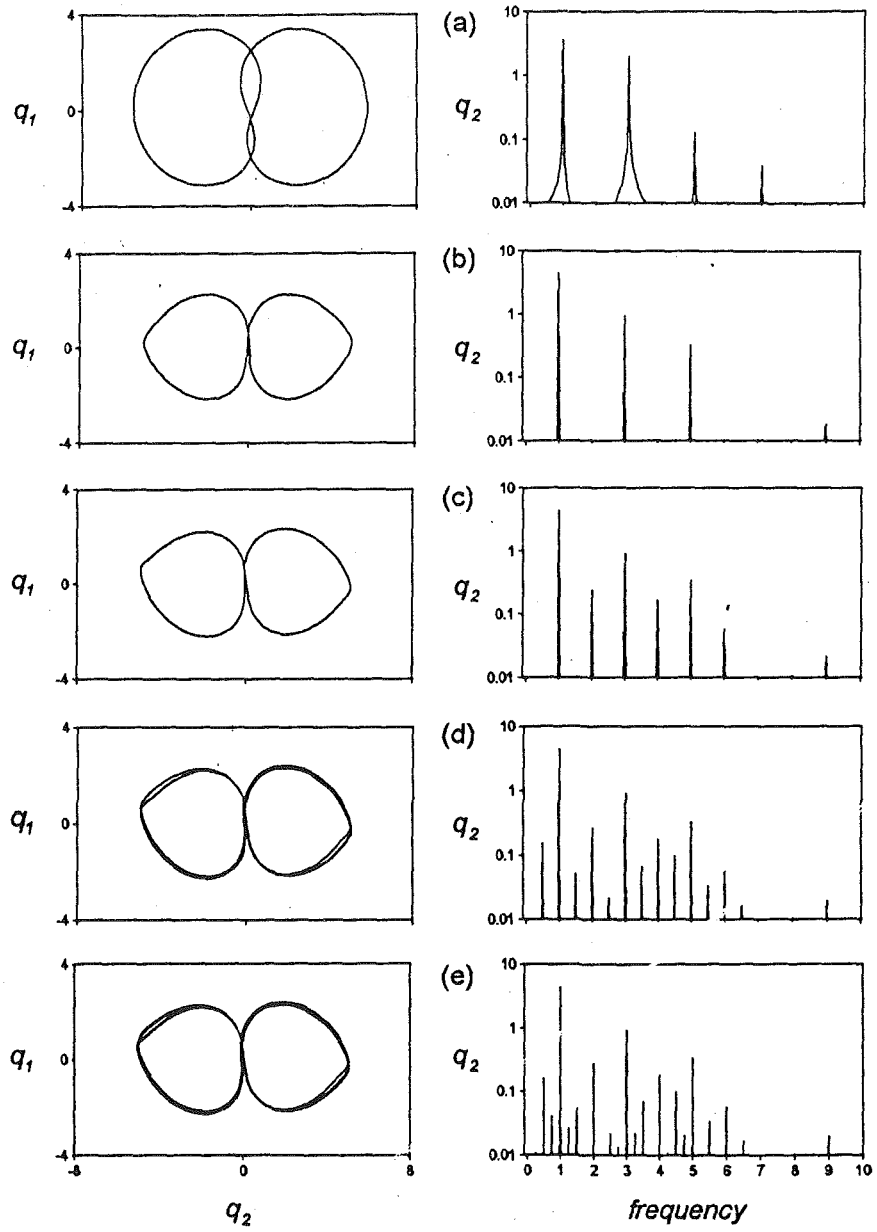


Fig. 6 Two-dimensional projections of the limit cycle onto the q_1–q_2 plane (left) and the corresponding FFT of q_2 (right) at σ = (*a*) −0.54697, (*b*) −1.0081, (*c*) −1.0569, (*d*) −1.0642, and (*e*) −1.0651 on branch II. They show a symmetry-breaking bifurcation followed by a sequence of period-doubling bifurcations.

attacking the continuous system. The dynamics involved here is detailed in the next section. However, in parts (b) and (d), a hardening-spring behavior similar to that in Figs. $1(b)$ and (d) is erroneously predicted by the two-mode discretization approach.

5 Dynamic Solutions

The frequency-response curves for the case of $n = 3$, as shown in Figs. $2(a)$ and (c), involve both single-mode and two-mode solutions. When σ decreases from a large value, say 100, or Ω decreases and approaches the natural frequency of the third mode, both types of solutions are possible. Hence, the response depends on the initial conditions. If the response is a single-mode solution, it remains stable as σ decreases until a subcritical (reverse) pitchfork bifurcation occurs at point A ($\sigma = 3.4698$), which leads to a jump to a stable two-mode equilibrium solution. The amplitude of this two-mode solution decreases as σ decreases further. It remains stable until a supercritical Hopf bifurcation occurs at point B ($\sigma = 1.1395$), which gives birth to stable small limit-cycle solutions. When σ increases from −100, the amplitude of the single-mode solution slowly increases. Once σ reaches point C ($\sigma \simeq -3.6669$), a saddle-node bifurcation takes place and leads to a jump to a far away solution, which is a chaotic attractor.

Figure 3 is a schematic bifurcation diagram for the orbits of the modulation Eqs. $(24)-(27)$ in the neighborhood of the unstable foci shown in Figs. $2(a)$ and (c). We calculate the limit cycles and their stability by using a combination of a shooting technique and Floquet theory (Nayfeh and Balachandran, 1995). Five branches of orbits are found.

On branch I, when σ decreases past the Hopf bifurcation point, the small limit cycle born at σ_1 increases in size, deforms, and then undergoes a sequence of period-doubling bifurcations, leading to chaos, as shown in Figs. $4(a)-(e)$. The responses come in pairs as solutions with q_2 and $-q_2$ coexist. As σ de-

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

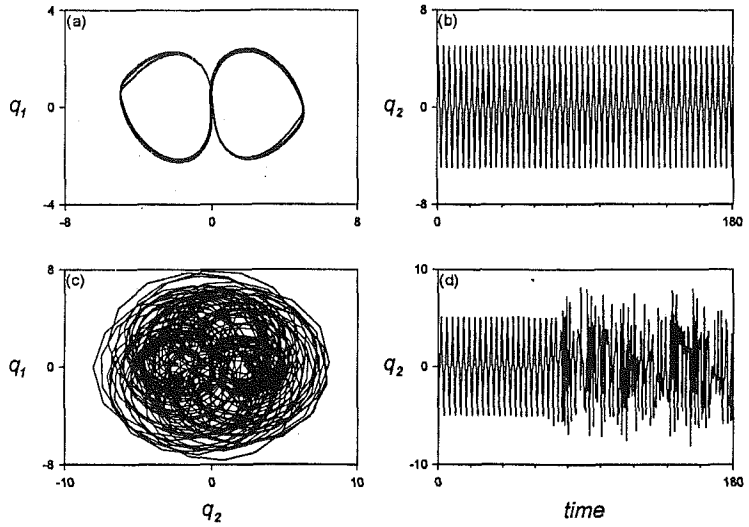


Fig. 7(*a*) A two-dimensional projection of the chaotic attractor onto the q_1–q_2 plane and (*b*) the time history of q_2 at $\sigma = -1.0660$. (*c*) A two-dimensional projection of the trajectory onto the q_1–q_2 plane and (*d*) the time history of q_2 at $\sigma = -1.0661$, where an explosive bifurcation occurs on branch II.

creases further, the local chaotic attractor encounters an explosive bifurcation (Thompson et al., 1994; Nayfeh and Balachandran, 1995) at $\sigma \simeq 0.71061$, as shown in Fig. 5. This type of interior crisis is due to a collision of the chaotic attractor with a regular saddle cycle. As a result, the chaotic attractor is abruptly enlarged to a chaotic attractor (on branch V), which includes the smaller one as a subset. Along the same branch, as σ decreases, a sequence of fragmented reverse period-doubling bifurcations occurs, leading to a singly winding limit cycle, as shown in Figs. $4(f)-(i)$. For example, a reverse p-4 bifurcation occurs at $\sigma \simeq 0.41815$, which leads to a stable period-two solution. It loses stability via a cyclic-fold bifurcation at $\sigma \simeq 0.41773$ and gives way to a chaotic attractor on branch V. A reverse p-2 bifurcation occurs at $\sigma \simeq -0.36348$, leading to a stable period-one solution. The period-one solution loses stability via a cyclic-fold bifurcation at $\sigma \simeq -0.36787$, resulting in a jump to a chaotic attractor on branch V.

On branch II, a cyclic-fold bifurcation occurs at $\sigma \simeq -0.54697$, leading to a chaotic attractor on branch V. As σ decreases, the symmetric limit cycle undergoes a symmetry-breaking bifurcation followed by a sequence of period-doubling bifurcations, culminating in chaos, as shown in Fig. 6. The ensuing chaotic attractor persists for a while before an explosive bifurcation occurs at $\sigma \simeq -1.0661$, as shown in Fig. 7. Again, it leads to a chaotic attractor on branch V.

On branch III, a cyclic-fold bifurcation takes place at $\sigma \simeq 3.93572$, resulting in a two-mode equilibrium solution. As σ decreases, the twin asymmetric limit cycles undergo a sequence of period-doubling bifurcations, as shown in Figs. $8(a)-(e)$. As σ decreases to 3.29825 approximately, the ensuing chaotic

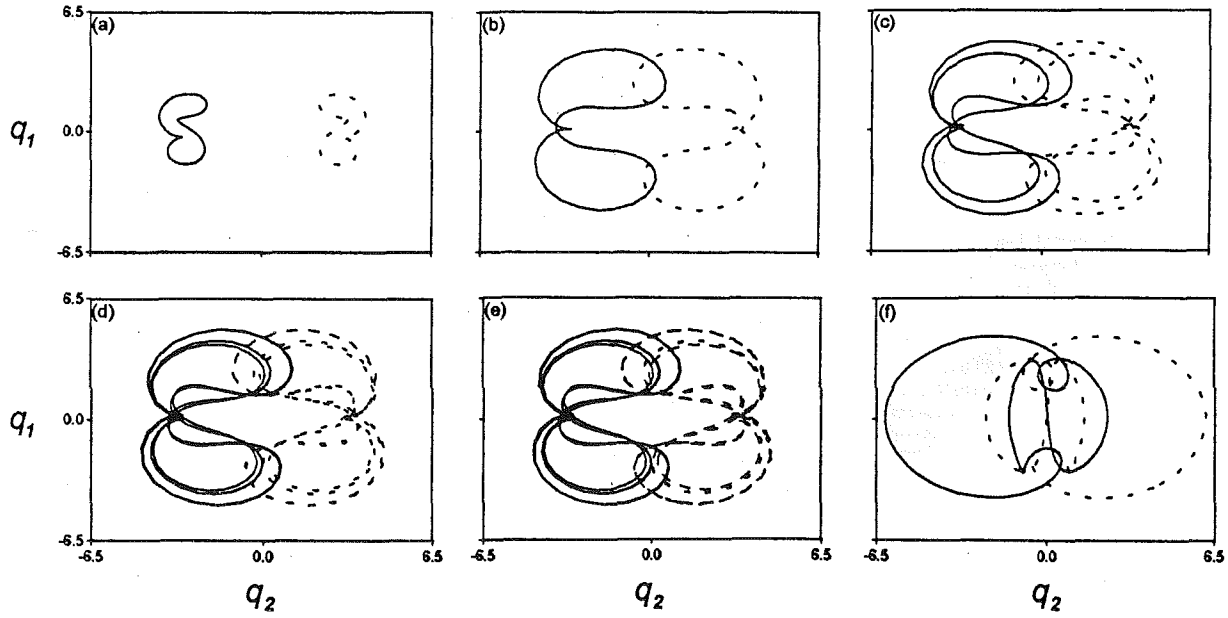


Fig. 8 Two-dimensional projections of the limit cycle onto the q_1–q_2 plane at σ = (*a*) 3.9357, (*b*) 3.3515, (*c*) 3.30577, (*d*) 3.29997, (*e*) 3.29971, and (*f*) 1.02983 on branch III. They show a sequence of period-doubling bifurcations from (*a*) to (*e*) and a reverse period-doubling bifurcation (*f*).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

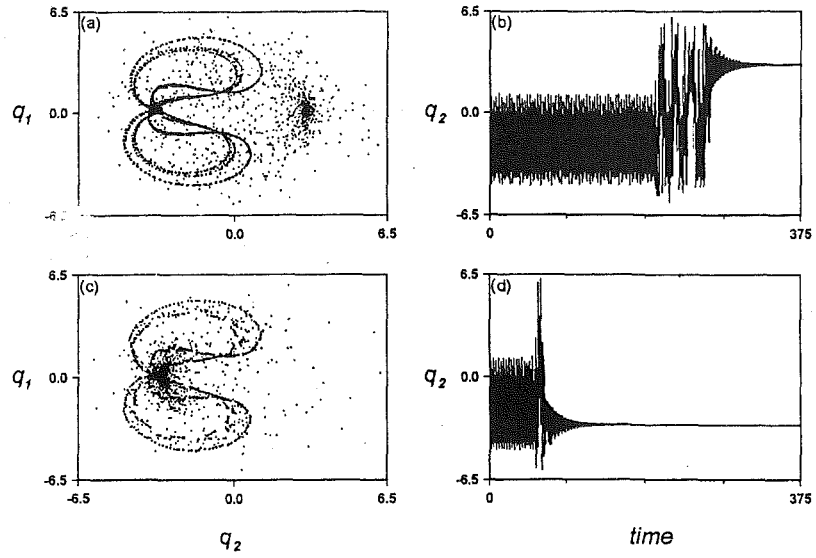


Fig. 9 Two-dimensional projections of the trajectory (discrete data points) onto the q_1–q_2 plane (left) and the time histories of q_2 (right) with the time steps (*a*, *b*) 0.125 and (*c*, *d*) 0.0625 used in the numerical simulations at $\sigma = 3.29825$, where a crisis occurs on branch III.

attractor collides with its basin boundary (boundary crisis) (Grebogi et al., 1983; Nayfeh and Balachandran, 1995), resulting in its destruction (Fig. 9). This type of instability leads to one of the two possible two-mode equilibrium solutions. The resulting attractor is quite sensitive to small disturbances or small variations in the initial conditions, as evident from Fig. 9 where two different time steps were used in the numerical integration. As σ decreases, the unstable period-one solution encounters a reverse period-doubling bifurcation at $\sigma \simeq 1.02983$ and gives way to a stable period-one limit cycle (Fig. 8(*f*)).

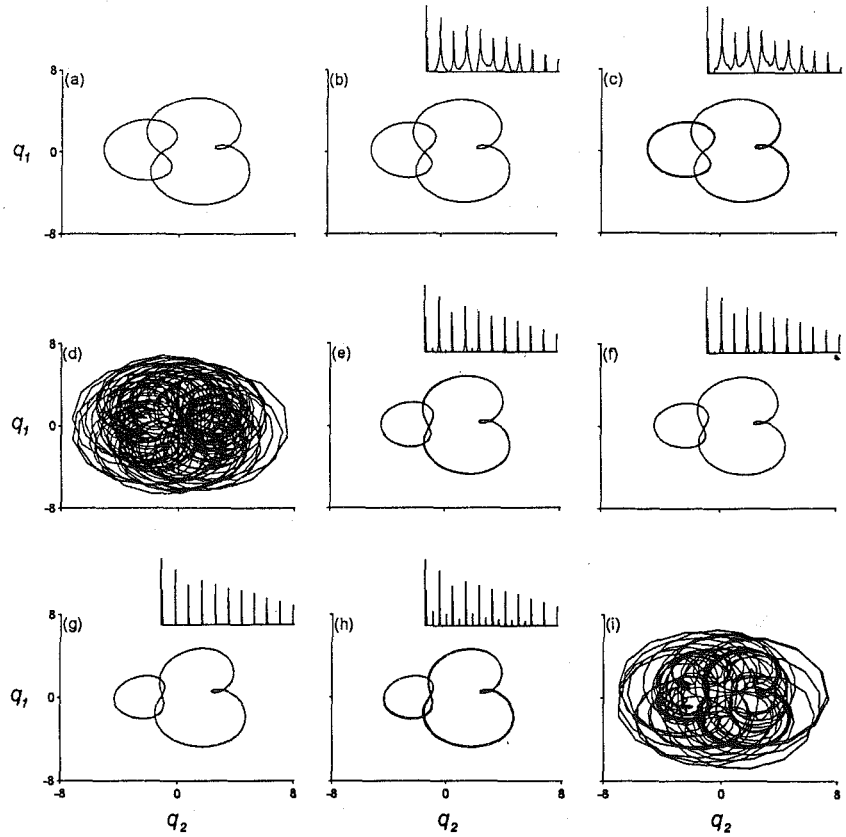


Fig. 10 Two-dimensional projections of the phase portraits onto the q_1–q_2 plane at $\sigma =$ (*a*) 0.52681 (*p*-1), (*b*) 0.50579 (*p*-1), (*c*) 0.50569 (*p*-2), (*d*) 0.31804 (chaos), (*e*) 0.31803 (*p*-2), (*f*) 0.318 (*p*-1), (*g*) 0.2303 (*p*-1), (*h*) 0.2261 (*p*-2), and (*i*) 0.2227 (chaos) on branch IV. They show three sequences of period-doubling bifurcations. The ensuing chaos encounters an explosive bifurcation and becomes a larger one (as shown in (*d*) and (*i*)).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

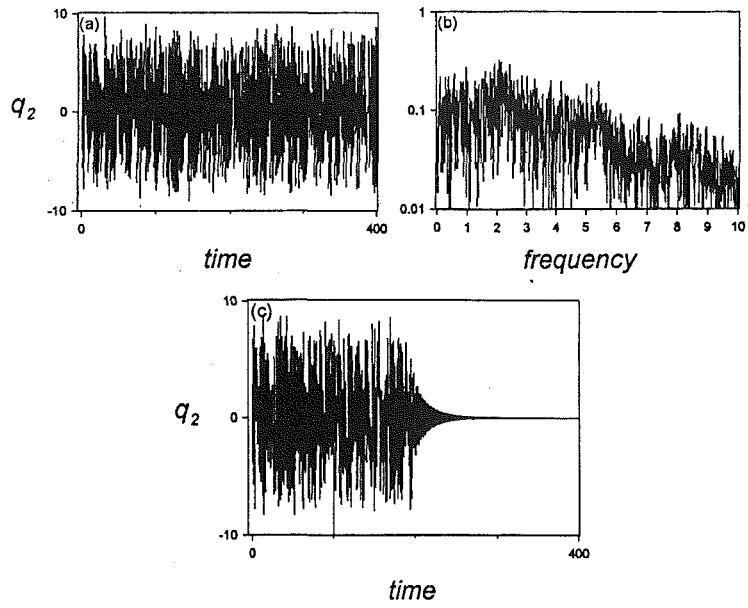


Fig. 11(*a*) The time history and (*b*) the FFT of q_2 at $\sigma = -3.66686$ on branch V. The chaotic attractor persists before a crisis occurs at $\sigma = -3.66687$, as shown in (c) the time history of q_2.

This orbit loses stability with one of the Floquet multipliers crossing the unit circle either through $+1$ as σ decreases or through -1 as σ increases. Both cases lead to a jump to a stable period-one solution on branch I.

On branch IV, a cyclic-fold bifurcation occurs at the right end; i.e., $\sigma \simeq 0.52681$. As σ increases, the limit cycle jumps to a chaotic attractor on branch V. As σ decreases, the twin asymmetric limit cycles undergo a sequence of period-doubling bifurcations, as shown in Figs. $10(a)-(c)$, leading to a chaotic attractor. The chaotic attractor suffers an explosive bifurcation like the one shown in Fig. $10(d)$. Along the branch of unstable period-one limit cycles, a reverse period-doubling bifurcation occurs at $\sigma \simeq 0.318$ and brings the unstable limit cycle back to a stable one. As σ decreases further, the twin asymmetric limit cycles undergo a sequence of period-doubling bifurcations again (Figs. $10(e)-(h)$), leading to a chaotic attractor. Again, this chaotic attractor suffers an explosive bifurcation, leading to its enlargement (Fig. $10(i)$).

As σ decreases, the large chaotic attractor on branch V remains stable until σ reaches -3.66687, which is close to the location where a saddle-node bifurcation of the single-mode equilibrium solution occurs. In Figs. $11(a)$ and (b), we show the time history and spectrum of the attractor at $\sigma \simeq -3.66686$. The broad spectrum of this attractor (Fig. $11(b)$) and its Lyapunov exponents (2.25, 0.0, -0.15, -2.40) confirm its chaotic nature. As σ decreases to -3.66687 approximately, a boundary crisis occurs, as shown in Fig. $11(c)$, which involves the collision of the chaotic attractor with its own basin boundary, resulting in the destruction of the attractor and its basin boundary and leading to a single-mode equilibrium solution.

6 Summary

We study the response of an externally excited infinitely long circular cylindrical shell with a one-to-one internal resonance to a primary excitation of one of its two orthogonal flexural modes. The obtained modulation equations reflect the symmetry property of the system. We compare the results derived by directly attacking the governing partial differential equations with those derived from a two-mode discretization approach. We show that the latter fail to account for the spatial variation at second order and completely miss the influence of the quadratic nonlinearity on the interaction coefficients.

We study the equilibrium solutions of the modulation equations and show that, for the second flexural modes, the frequency-response curve exhibits a hardening-spring behavior for both approaches. However, the response amplitude is underestimated and a pitchfork bifurcation is erroneously predicted with the discretized approach. For the third flexural modes, the direct approach predicts a softening-spring behavior in the frequency-response curves, whereas the discretization approach predicts a hardening-spring behavior.

The complexities of the two-mode equilibrium solution below the Hopf bifurcation frequency are explored. These include bubble structures (the occurrence of cascades of period-doubling and reverse period-doubling bifurcations) (Knobloch and Weiss, 1983; Nayfeh and Balachandran, 1995), the coexistence of multiple attractors, cyclic-fold and symmetry-breaking bifurcations, explosive bifurcations, and boundary crises. The explosive bifurcations are due to the collision of two local chaotic attractors with a regular saddle cycle, resulting in the merging of the chaotic attractors and the production of a single chaotic attractor that is larger than them. To predict such bifurcations, one needs to trace the local growth of chaotic attractors and their proximity to the saddle limit cycle.

Acknowledgment

This work was supported by the Air Force Office of Scientific Research under Grant No. F49620-92-J-0197.

References

Chen, J. C., and Babcock, C. D., 1975, "Nonlinear Vibrations of Cylindrical Shells," *AIAA Journal,* Vol. 13, pp. 868–876.

Evensen, D. A., 1966, "Nonlinear Flexural Vibrations of Thin Circular Rings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 553–560.

Grebogi, C., Ott, E., and Yorke, J. A., 1983, "Crises, Sudden Changes in Chaotic Attractors, and Transient Chaos," *Physica D,* Vol. 7, pp. 181–200.

Knobloch, E., and Weiss, N. O., 1983, "Bifurcations in a Model of Magnetoconvection," *Physics of Fluids,* Vol. 9, pp. 379–407.

Maewal, A., 1981, "Nonlinear Harmonic Oscillations of Gyroscopic Structural Systems and the Case of a Rotating Ring," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 627–633.

Maewal, A., 1986, "Miles' Evolution Equations for Axisymmetric Shells: Simple Strange Attractors in Structural Dynamics," *International Journal of Nonlinear Mechanics,* Vol. 21, pp. 433–438.

Maganty, S. P., and Bickford, W. B., 1987, "Large Amplitude Oscillations of Thin Circular Rings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 315–322.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Maganty, S. P., and Bickford, W. B., 1988, "Influence of Internal Resonance on the Non-Linear Oscillations of a Circular Ring under Primary Resonance Conditions," *Journal of Sound and Vibration,* Vol. 122, pp. 521–547.

McIvor, I. K., 1966, "The Elastic Cylindrical Shell under Radial Impulse," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 831–837.

Miles, J. W., 1984, "Resonantly Forced Surface Waves in a Circular Cylinder," *Journal of Fluid Mechanics,* Vol. 149, pp. 15–31.

Nayfeh, A. H., 1973, *Perturbation Methods,* John Wiley and Sons, New York.

Nayfeh, A. H., 1981, *Introduction to Perturbation Techniques,* John Wiley and Sons, New York.

Nayfeh, A. H., and Balachandran, B., 1995, *Applied Nonlinear Dynamics,* John Wiley and Sons, New York.

Raouf, R. A., and Nayfeh, A. H., 1990, "One-to-One Autoparametric Resonances in Infinitely Long Cylindrical Shells," *Computers & Structures,* Vol. 35, pp. 163–173.

Thompson, J. M. T., Stewart, H. B., and Ueda, Y., 1994, "Safe, Explosive, and Dangerous Bifurcations in Dissipative Dynamical Systems," *Physical Review E,* pp. 1019–1027.

APPENDIX

$$\omega_n^4 - [n^2 + 1 + \alpha^2(n^2-1)^2]\omega_n^2 + \alpha^2 n^2 (n^2-1)^2 = 0, \quad n = 0, 1, 2, \ldots \tag{A1}$$

$$\Gamma_n = \frac{1}{n}[1 + \alpha^2(n^2-1)^2 - \omega_n^2] \tag{A2}$$

$$\Delta_n = 4(n^2 - \omega_n^2)[1 - 4\omega_n^2 + \alpha^2(4n^2-1)^2] - 4n^2 \tag{A3}$$

$$\Theta_n = 4n^2\alpha^2(4n^2-1)^2 \tag{A4}$$

$$c_1 = \tfrac{1}{2}\Gamma_n^2(n^2 - \omega_n^2) - \Gamma_n(n^3 + n) + \tfrac{3}{4}n^2 + \alpha^2(9\Gamma_n n^5 - \tfrac{9}{2}\Gamma_n n^3 - 9n^4 + \tfrac{33}{4}n^2 - \tfrac{3}{2}) \tag{A5}$$

$$c_2 = \tfrac{1}{2}\Gamma_n^2(n^2 + \omega_n^2) - \Gamma_n n + \tfrac{1}{4}n^2 + \alpha^2(-\tfrac{1}{2}\Gamma_n n^3 - n^4 + \tfrac{11}{4}n^2 - \tfrac{3}{2}) \tag{A6}$$

$$c_3 = -\Gamma_n^2(n^2 + \omega_n^2) + \Gamma_n(2n^3 + 2n) - \tfrac{3}{2}n^2 + \alpha^2(-18\Gamma_n n^5 + 9\Gamma_n n^3 + 18n^4 - \tfrac{33}{2}n^2 + 3) \tag{A7}$$

$$c_4 = -\tfrac{1}{2}\Gamma_n^2(n^2 - \omega_n^2) + \Gamma_n n - \tfrac{1}{4}n^2 + \alpha^2(\tfrac{1}{2}\Gamma_n n^3 + n^4 - \tfrac{11}{4}n^2 + \tfrac{3}{2}) \tag{A8}$$

$$d_1 = 2\Gamma_n(n^2 - \omega_n^2) + \tfrac{1}{2}n^3 - n \tag{A9}$$

$$d_2 = -4\Gamma_n n^2 - n^3 + 2n \tag{A10}$$

$$\begin{aligned}
8\omega_n(1+\Gamma_n^2)\Lambda_1 = {} & \frac{1}{2}\Gamma_n^2\omega_n^2 + \frac{3}{2}\Gamma_n^2 n^4 - \frac{9}{2}\Gamma_n^2 n^2 - \frac{3}{2}\Gamma_n n^3 - \frac{9}{8}n^4 + \frac{3}{2}n^2 + \frac{1}{\Delta_n}[4\Gamma_n^2\omega_n^4 c_1 - 2\Gamma_n^2\omega_n^2 n d_1 + 4\Gamma_n^2 n^4 c_1 \\
& + 2\Gamma_n^2 n^3 d_1 + 4\Gamma_n\alpha^2 n^2 d_1 - 64\Gamma_n\alpha^2\omega_n^2 n^4 d_1 - 76\alpha^2 n^5 d_1 - 8\Gamma_n n^5 c_1 + 36\Gamma_n\alpha^2\omega_n^2 n^3 c_1 + 32\Gamma_n\alpha^2\omega_n^2 n^2 d_1 \\
& - 4\Gamma_n\alpha^2\omega_n^2 d_1 - 2n d_1 - 50\Gamma_n\alpha^2 n^4 d_1 + 16\Gamma_n\omega_n^4 d_1 + 8\Gamma_n\omega_n^2 n^3 c_1 - 16\Gamma_n\omega_n^2 n^2 d_1 + 4n^3 d_1 - 4n^2 c_1 \\
& - 4\Gamma_n\omega_n^2 d_1 - 4\Gamma_n n^4 d_1 + 72\alpha^2\omega_n^2 n^4 c_1 - 36\Gamma_n\alpha^2 n^5 c_1 - 72\alpha^2 n^6 c_1 - 66\alpha^2\omega_n^2 n^2 c_1 + 12\alpha^2\omega_n^2 c_1 + 66\alpha^2 n^4 c_1 \\
& + 50\alpha^2 n^3 d_1 - 12\alpha^2 n^2 c_1 - 8\alpha^2 n d_1 - 4\omega_n^2 n^3 d_1 - 6\omega_n^2 n^2 c_1 + 8\omega_n^2 n d_1 - 8\Gamma_n^2\omega_n^2 n^2 c_1 + 100\Gamma_n\alpha^2 n^6 d_1 \\
& + 16\alpha^2 n^7 d_1 + 72\Gamma_n\alpha^2 n^7 c_1 - 72\Gamma_n\alpha^2\omega_n^2 n^5 c_1 + 8n^4 c_1] + \frac{1}{\Theta_n}[50\Gamma_n\alpha^2 n^4 d_2 + 12\alpha^2 n^2 c_3 - 50\alpha^2 n^3 d_2 \\
& - 4\Gamma_n\alpha^2 n^2 d_2 - 8n^4 c_3 + 4n^2 c_3 - 4n^3 d_2 + 4\Gamma_n n^4 d_2 - 66\alpha^2 n^4 c_3 + 8\alpha^2 n d_2 + 2n d_2 - 4\Gamma_n^2\omega_n^2 n^2 c_3 \\
& - 2\Gamma_n^2\omega_n^2 n d_2 - 4\Gamma_n^2 n^4 c_3 - 2\Gamma_n^2 n^3 d_2 + 76\alpha^2 n^5 d_2 + 8\Gamma_n n^5 c_3 + 36\Gamma_n\alpha^2 n^5 c_3 + 72\alpha^2 n^6 c_3 - 100\Gamma_n\alpha^2 n^6 d_2 \\
& - 16\alpha^2 n^7 d_2 - 72\Gamma_n\alpha^2 n^7 c_3] + \frac{c_2}{1 - 4\omega_n^2 + \alpha^2}[-2\Gamma_n\alpha^2 n^3 + n^2 - 4\Gamma_n n - 4\alpha^2 n^4 + 11\alpha^2 n^2 - 6\alpha^2 \\
& + 2\Gamma_n^2\omega_n^2 + 2\Gamma_n^2 n^2] + \frac{c_4}{1 + \alpha^2}[4\Gamma_n\alpha^2 n^3 - 2n^2 + 8\Gamma_n n + 8\alpha^2 n^4 - 22\alpha^2 n^2 + 12\alpha^2 - 4\Gamma_n^2 n^2 + 4\Gamma_n^2\omega_n^2]
\end{aligned} \tag{A11}$$

$$\begin{aligned}
8\omega_n(1+\Gamma_n^2)\Lambda_2 = {} & -\frac{5}{2}\Gamma_n^2\omega_n^2 + \frac{1}{2}\Gamma_n^2 n^4 - \frac{3}{2}\Gamma_n^2 n^2 - \frac{1}{2}\Gamma_n n^3 - \frac{3}{8}n^4 + \frac{1}{2}n^2 + \frac{1}{\Delta_n}[64\Gamma_n\alpha^2\omega_n^2 n^4 d_1 - 36\Gamma_n\alpha^2\omega_n^2 n^3 c_1 \\
& - 32\Gamma_n\alpha^2\omega_n^2 n^2 d_1 + 4\Gamma_n\alpha^2\omega_n^2 d_1 + 50\Gamma_n\alpha^2 n^4 d_1 - 4\Gamma_n\alpha^2 n^2 d_1 - 16\Gamma_n\omega_n^4 d_1 - 8\Gamma_n\omega_n^2 n^3 c_1 + 16\Gamma_n\omega_n^2 n^2 d_1 \\
& + 4\Gamma_n\omega_n^2 d_1 + 4\Gamma_n n^4 d_1 - 72\alpha^2\omega_n^2 n^4 c_1 + 66\alpha^2\omega_n^2 n^2 c_1 - 12\alpha^2\omega_n^2 c_1 - 66\alpha^2 n^4 c_1 - 50\alpha^2 n^3 d_1 + 8\alpha^2 n d_1 \\
& + 4\omega_n^2 n^3 d_1 + 2n d_1 + 6\omega_n^2 n^2 c_1 - 8\omega_n^2 n d_1 - 8n^4 c_1 - 4n^3 d_1 + 4n^2 c_1 + 12\alpha^2 n^2 c_1 + 2\Gamma_n^2\omega_n^2 n d_1 \\
& - 4\Gamma_n^2\omega_n^4 c_1 - 4\Gamma_n^2 n^4 c_1 - 2\Gamma_n^2 n^3 d_1 + 76\alpha^2 n^5 d_1 + 8\Gamma_n n^5 c_1 + 36\Gamma_n\alpha^2 n^5 c_1 + 72\alpha^2 n^6 c_1 - 100\,\Gamma_n\alpha^2 n^6 d_1 \\
& - 16\alpha^2 n^7 d_1 - 72\Gamma_n\alpha^2 n^7 c_1 + 8\Gamma_n^2\omega_n^2 n^2 c_1 + 72\Gamma_n\alpha^2\omega_n^2 n^5 c_1] + \frac{1}{\Theta_n}[4n^2 c_3 + 50\Gamma_n\alpha^2 n^4 d_2 - 4\Gamma_n\alpha^2 n^2 d_2 \\
& + 4\Gamma_n n^4 d_2 - 66\alpha^2 n^4 c_3 - 50\alpha^2 n^3 d_2 + 2n d_2 + 12\alpha^2 n^2 c_3 + 8\alpha^2 n d_2 - 8n^4 c_3 - 4n^3 d_2 - 4\Gamma_n^2\omega_n^2 n^2 c_3 \\
& - 2\Gamma_n^2\omega_n^2 n d_2 - 4\Gamma_n^2 n^4 c_3 - 2\Gamma_n^2 n^3 d_2 + 76\alpha^2 n^5 d_2 + 8\Gamma_n n^5 c_3 + 36\Gamma_n\alpha^2 n^5 c_3 + 72\alpha^2 n^6 c_3 - 100\,\Gamma_n\alpha^2 n^6 d_2 \\
& - 16\alpha^2 n^7 d_2 - 72\Gamma_n\alpha^2 n^7 c_3] + \frac{c_2}{1 - 4\omega_n^2 + \alpha^2}[-2\Gamma_n\alpha^2 n^3 + n^2 - 4\Gamma_n n - 4\alpha^2 n^4 \\
& + 11\alpha^2 n^2 - 6\alpha^2 + 2\Gamma_n^2\omega_n^2 + 2\Gamma_n^2 n^2]
\end{aligned} \tag{A12}$$

$$2\omega_n(1 + \Gamma_n^2)F = f \tag{A13}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Transient Response in a One-Dimensional Model of Thermoelastic Contact

Z. S. Olesiak

Yu. A. Pyryev[1]

Department of Applied Mathematics and Mechanics, University of Warsaw, 02-097 Warsaw, Poland

We consider two layers of different materials with the initial gap between them in the field of temperature with imperfect boundary conditions in Barber's sense. The model we discuss is that of two contacting rods (Barber and Zhang, 1988) which in the case of a single rod was devised and discussed by Dundurs and Comninou (1976, 1979). In this paper we try to make a step further in the investigation of the essentially nonlinear problem. Though we consider a system of the linear equations of thermoelasticity the nonlinearity is induced by the boundary conditions dependent on the solution. We present an algorithm for solving the system of equations based on Laplace's transform technique. The method of solution can be used also in the dynamical problems with inertial terms taken into account. The numerical results have been obtained by a kind of computational simulation.

1 Introduction

The surface forces, or tractions can be transmitted only by contact of the surfaces, or their parts, bounding the solids. In general case the regions of contact are not known beforehand and can be established from the solutions of the pertinent boundary value problems, in the simplest case from those of the linear theory of elasticity of two contacting solids. Historically we can distinguish two kinds of contact problems. The so-called Boussinesq's, or the first contact problem (1885), posed and solved by W. Kelvin already in 1848, consisted in finding the distribution of stresses generated by known concentrated force applied at a point of the boundary of an elastic semi-space. The main point of interest was to find how the stresses spread out from a point of application of the force (''the stress propagation''). In the case when the value, or the distribution of tractions is not known, when two solids are in contact over a certain regular surface, we deal with the second contact problem, stemming from the well-known Hertz's concept (1881). From the standpoint of mathematics, the problem of the ideal contact can be reduced to solving a type of mixed boundary value problem. In the case a contact with a rigid solid it was found that in the distribution of stress tensor components there may exist singularities, dependent on the jumps in the values of the displacement vector derivatives. For many years the investigations in this field were being concentrated on solving problems with different geometry and for different models of solids like elastic, elastic-plastic, the influence of viscous effects, porosity, etc. It seems that the paper by George and Sneddon (1962) was the first one in which the effect of punch heating the semispace and the field of temperature was taken into account. Then it was believed that the contact problems of thermoelasticity could be reduced to a little bit more complicated mixed boundary value problems and the difficulty layed in the fact that we dealt with a system of equations of thermoelasticity instead of Navier's partial differential equation for an elastic solid or solids in contact. The contact problems of elasticity, at least many of them, could be reduced to a superposition of a problem of isothermal elasticity and a simple auxiliary one of thermoelasticity (see Olesiak, 1993). The method of superposition can be applied for functions as well as distributions, also in the cases when singularities occur (the corresponding statement in Barber and Comninou (1989) p. 34 is not precise). Next it was discovered that to assume the smoothness of the contacting surfaces is not always sufficient to obtain a solution of physical significance. The solutions of the mathematical boundary value problems of thermoelasticity were not necessarily those of the physical meaning, namely the tensile contact tractions were obtained when the surface was not convex. For concave surfaces it was necessary to assume either two-sided constraints or to consider inequalities for unknown regions of contact and single-sided constraints. In the papers of J. Dundurs, Maria Comninou, and others a new class of the contact problems with wavy surfaces was considered and developed.

It was J. R. Barber (1978) who discovered that the solution of the classical problem of thermoelasticity for a cooled sphere was leading to an abnormal behavior and nonphysical tensile tractions. Though the indentation into a semi-space of a heated sphere and a cooled sphere are alike from the standpoint of the mixed boundary value problems, in the first case the convex surface of the sphere is in contact with the convex surface of the deformed bounding plane of the semi-space. In the case of a cooled sphere its surface remains convex while the deformed surface of the plane bounding the semi-space due to cooling becomes concave. The condition that the nontensile inequality depends on the direction (rather the sense) of the heat flux into the body of higher distortivity (see Barber and Comninou (1989) p. 7) is not necessarily true in the presence of other fields like that of mass diffusion or mechanical body forces. To explain the phenomenon that the tensile tractions may occur Barber has developed the rod model, proposed in a paper by Dundurs and Comninou (1976). We shall not repeat presenting the idea of the model. It has been discussed a couple of times and presumably most fully presented in a paper by Barber and Zhang (1988). The model is by assumption a statical one, the results, however, and the drawn conclusions supply the time-dependent behavior. In this paper we have made an attempt to verify the results obtained by Barber and Zhang by taking as our point of departure a time-dependent system of elliptic-parabolic partial differential equations of thermoelasticity with boundary conditions and initial conditions nonlinearly coupled through the solutions.

[1] Permanent address: Chair of mechanics, University of Lviv, Lviv, Ukraine.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, June 27, 1994; final revision, Mar. 12, 1995. Associate Technical Editor: F. Y. M. Wan.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

There is a certain missunderstanding of the terminology used by mathematicians and in applications of mathematics in the problems of contact mechanics. The solutions of the classical contact problems of thermoelasticity usually possess unique solutions and the sense of the thermal flux vector, or the occurrence of the tensile tractions, has nothing to do with the uniqueness of the solution. Extra inequalities of physical meaning can be either satisfied or not. Nonuniqueness of the solutions of the mixed boundary value problems would appear if we were seeking the solutions of the pertinent problems in the classes of generalized functions (compare A. C. McBride, *Fractional Calculus and Integral Transforms of Generalized Functions,* Pitman Publ., Inc., 1979). However, also in the cases considered by McBride the solutions become classical and unique when one assumes that the energy density has to be finite. The fact that the solution for the contact problem of a cooled sphere, or alike, has no physical significance on a part of boundary means that the assumed model of an elastic solid with classical boundary conditions is not suitable in such cases. Presumably the first problem of the contact mechanics (isothermal) in which the nonphysical solutions occurred was that considered by V. M. Abramov already in 1937 (Doklady AN SSSR, Vol. 17). Abramov envisaged an isothermal contact of a punch with an elastic semi-space with a rough bounding plane and obtained a nonphysical solution with overlapping regions in the neighborhood of the singularities.

Later on, in the 1970s, similar problems were being discussed for problems of cracks in the bonding plane of two dissimilar solids. In the cases when the theoretical solution has no physical meaning it is necessary to construct another mathematical model. In the case of contact problems of thermoelasticity the conditions of imperfect contact has been conceived (see Comninou and Dundurs, 1979; Dundurs and Comninou, 1976; Barber and Comninou, 1989) connected with the notion of interfacial thermal resistance. It was of little interest to find a special case of the contact resistance such that the solution of the system of differential equation of thermoelasticity was still unique (Duvaut, 1976). The problems with multiple solutions of the equations of the classical mechanics are known. In the theory of the dynamical systems one encounters multiple solutions, fixed points of functions, points of bifurcations, conditions of stability, the notion of cycles, etc., everything for the systems of the ordinary, nonlinear differential equations. The fixed points which we discuss in this paper correspond to the "cycle S^2" in the theory of dynamical systems. It was unexpected to find the behavior known from the solutions of certain nonlinear differential equations in the case of a set of the linear partial (of elliptic-parabolic type) differential equations of thermoelasticity.

The purpose of the paper is to investigate the time behavior of the contact characteristics since the beginning of the wall temperature change to the exit of the system on the steady-state regime. We have applied more rigorous mathematical methods to the question of existence and uniqueness of the steady-state solutions as compared with those of Barber and Zhang (1988). The conditions for the uniqueness and nonuniqueness of the steady-state solution have been given in the analytical form for a special case of the thermal resistance. The effect of thermal history on the final (and intermediate) states in cases of nonuniqueness are explored more fully, using the situations where the wall temperatures approach their final condition in various ways from the different initial states. Barber and Zhang (1988) considered only those transient processes resulting from the perturbation of an unstable steady-state, with constant wall temperatures.

2 Basic Equations

For two solids in an ideal contact we have the following boundary conditions of the thermal contact over a region Ω (Carslaw and Jaeger, 1959)

$$\theta_1(\xi, 0, t) = \theta_2(\xi, 0, t),$$

$$-K_1 \frac{\partial \theta_1(\xi, 0, t)}{\partial x} = -K_2 \frac{\partial \theta_2(\xi, 0, t)}{\partial x},$$

$$\text{for:} \quad \xi \in \Omega, \quad t \in (0, \infty). \quad (1)$$

Here θ_1 and θ_2 denote the temperatures of the first and the second solid, with respect to the temperature of the initial state, K_1 and K_2 are the coefficients of thermal conductivity, respectively. However, in practice the contact between two solids, as a rule, is not perfect; there may exist the gaps filled with a gas (for example air), thin layers of lubricant, and water. The heat is no longer transmitted either through the conduction or linear heat radiation, and consequently the boundary thermal conditions cannot be modelled as the ideal contact of solids. Besides, the solids change in time their shapes and volumes due to the action of forces and heat fluxes, the surface bounding the solids is assumed to be rough. Thus we obtain the following conditions much better describing the physical reality:

$$\begin{cases} K_1 \dfrac{\partial \theta_1(\xi, 0, t)}{\partial x} - K_2 \dfrac{\partial \theta_2(\xi, 0, t)}{\partial x} = 0, \\ -K_1 \dfrac{\partial \theta_1(\xi, 0, t)}{\partial x} = \dfrac{\theta_1(\xi, 0, t) - \theta_2(\xi, 0, t)}{R}, \\ \xi \in \Omega, \quad t \in (0, \infty), \end{cases} \quad (2)$$

where R is the thermal resistance

$$R = R_{sep} H(g) + R_{con} H(p),$$

$$\Omega = \Omega_{con} \cup \Omega_{sep}, \; \partial\Omega_{con} \cap \partial\Omega_{sep} = l : g(\xi, t) = 0,$$

$$t_{sep} \cup t_{con} \in (0, \infty),$$

$$t \in t_{con}, \quad \xi \in \Omega_{con} \Leftrightarrow p(\xi, t) > 0,$$

$$t \in t_{sep}, \quad \xi \in \Omega_{sep} \Leftrightarrow g(\xi, t) > 0.$$

$H(\cdot)$ denotes the Heaviside's step distribution, while

$$\begin{cases} \sigma_{xx}^{(n)}(\xi, 0, t) = 0, \quad \sigma_{xz}^{(n)}(\xi, 0, t) = 0, \quad \sigma_{xy}^{(n)}(\xi, 0, t) = 0, \\ \xi \in \Omega_{sep}, \quad t \in t_{sep}, \quad n = 1, 2, \\ u_1(\xi, 0, t) = u_2(\xi, 0, t), \\ g(\xi, t) = 0, \quad \sigma_{xx}^{(1)}(\xi, 0, t) = \sigma_{xx}^{(2)}(\xi, 0, t) = -p(\xi, t), \\ \sigma_{xz}^{(n)}(\xi, 0, t) = 0, \quad \sigma_{xy}^{(n)}(\xi, 0, t) = 0, \\ \xi \in \Omega_{con}, \quad t \in t_{con}. \end{cases} \quad (3)$$

The conditions (2)–(3) could be made even more general if we had taken into account the friction and the possibility of tangential motion.

As the point of departure of our discussion we take the system of equations of coupled thermoelasticity (Nowacki, 1986), namely the generalized Navier equation:

$$\nabla^2 \mathbf{u} + \frac{1}{1 - 2\nu} \operatorname{grad} \operatorname{div} \mathbf{u} = \frac{\rho}{\mu} \ddot{\mathbf{u}} + \frac{2(1 + \nu)}{1 - 2\nu} \alpha \operatorname{grad} \theta, \quad (4)$$

and the thermal conductivity equation:

$$\nabla^2 \theta - \frac{1}{k} \frac{\partial \theta}{\partial t} - \eta \operatorname{div} \dot{\mathbf{u}} = -\frac{Q}{K} \quad (5)$$

where Q denotes a heat source, $\eta = \gamma \theta_0 / K$. In order to construct a suitable model one has to take into account the dependance on the temperature of solids and of thermal resistance. The problem states how the thermal resistance varies on the boundary $\Omega = \Omega_{con} \cup \Omega_{sep}$.

We can look upon the thermal resistance R as such a function, the regulating resistance, that collects the information from the

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

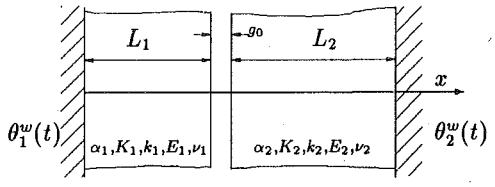


Fig. 1 Two layers heated from the walls

thermally contacting surfaces and supplies the law of the action of thermal fluxes. In this connection we can cite Stephan-Boltzmann's model of heat exchange (compare Carslaw and Jaeger, 1959). Such a piece of information can be collected from the entire regions of the solids in contacts or from a number of separate points of the solids. As an example we may consider a nonstationary system with automatic thermal control. In practice the thermal systems are self-controlling and the problem consists in finding the laws of the self-control. One of the characteristics of control is the thermal displacement of the surfaces $u_1(\xi, 0, t)$, $u_2(\xi, 0, t)$, $\xi \in \Omega$ in the case of gaps and the contact pressure $p(\xi, t)$ for contact regions. It is intuitively clear that the thermal resistance grows with the distance between solids $g(\xi, t)$, $\xi \in \Omega_{\text{sep}}$ in noncontact regions while for the greater contact pressure $p(\xi, t)$ it decreases. It is assumed that this function is continuous also at the transition from the gap to the contact regions. Though it has not been verified to which extent this assumption is necessary. A number of experimental investigations have been devoted to establish the dependance of the thermal resistance on the contact pressure and the gap between the solids outside the contact region (Thomas and Probert, 1970).

In this paper we prefer to consider a one-dimensional model of nonstationary thermal contact of two layers shown in Fig. 1. There is no practical difference in geometry as compared with the model considered in (Barber and Zhang, 1988) except that here we distinguish the boundary conditions in the direction perpendicular to x-axis. The model of two rods corresponds to the case when $\sigma_{yy} = \sigma_{zz} = 0$.

We denote the temperature $\theta(\xi, x, t) \equiv \theta(x, t)$ and the displacement vector $\mathbf{u} = (u, v, w)$, $u(\xi, x, t) \equiv u(x, t)$. Two layers ($n = 1, 2$) of the width L_n are placed between two rigid walls. We denote Young's moduli E_n, Poisson coefficients ν_n, coefficients of linear thermal expansion α_n, coefficients of heat conductivity K_n, and thermal diffusivity k_n. The walls conduct heat and the temperature of the walls changes in time $\theta_n^w(t) = \theta_0 + \vartheta_n\phi_n(t)$, $\phi_n(t) = 0$, $t < 0$. It is assumed that there is a small initial gap between the two layers $g_0 > 0$, $g_0/L_n \ll 1$ for $t < 0$. On the boundary we assume the condition of thermal resistance, i.e., a known function of distance between the solids outside the mechanical contact and a function of contact pressure (compare Barber, Dundurs, and Comninou (1979), Srinivasan and France (1985), and Thomas and Probert (1970)). The function of the thermal resistance can be represented in the following way:

$$R = R_{\text{sep}}(g)H(g) + R_{\text{con}}(p)H(p), \quad R_{\text{sep}}(g) \to \infty,$$
$$\text{for} \quad g \to \infty, \quad R_{\text{con}}(p) \to 0, \quad \text{for} \quad p \to \infty \quad (6)$$

where $H(.)$ denotes the Heaviside unit distribution. In the considered one-dimensional case we obtain the system of equations for $n = 1, 2$:

$$\frac{\partial}{\partial x}\left(\frac{\partial u_n(x, t)}{\partial x} - \frac{1 + \nu_n}{1 - \nu_n}\alpha_n\theta_n(x, t)\right) = 0, \quad (7)$$

$$\frac{\partial^2\theta_n(x, t)}{\partial x^2} - \frac{1}{k_n}\frac{\partial\theta_n(x, t)}{\partial t} - \eta_n\frac{\partial^2 u_n(x, t)}{\partial x\partial t} = 0, \quad t \in (0, \infty),$$
$$\text{for} \quad n = 1, \quad x \in (0, L_1), \quad \text{for} \quad n = 2,$$
$$x \in (L_1 + g_0, L_1 + g_0 + L_2). \quad (8)$$

Since the process is assumed to be slow we have neglected the inertia terms. It is also assumed that there are no heat sources within the two layers. The Duhamel relations

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\epsilon} + (\lambda\,\text{tr}\,\boldsymbol{\epsilon} - \gamma\theta)\mathbf{1}, \quad (9)$$

where $\gamma = (3\lambda + 2\mu)\alpha$, reduce to the following equations, for the layers $n = 1, 2$:

$$\sigma_{xx}^{(n)} = \frac{E_n}{1 - 2\nu_n}\left(\frac{1 - \nu_n}{1 + \nu_n}\frac{\partial u_n}{\partial x} - \alpha_n\theta_n\right)$$
$$\text{valid for} \quad \epsilon_{yy}^{(n)} = \epsilon_{zz}^{(n)} = 0, \quad (10)$$

and

$$\sigma_{xx}^{(n)} = E_n\left(\frac{\partial u_n}{\partial x} - \alpha_n\theta_n\right), \quad \text{valid for} \quad \sigma_{yy}^{(n)} = \sigma_{zz}^{(n)} = 0.$$

The last equation is valid likewise for two rods in contact (Barber and Zhang (1988)).

We arrive at the following mechanical boundary conditions. In the case when there exists a gap we obtain

$$u_1(0, t) = 0, \quad u_2(L_1 + g_0 + L_2, t) = 0,$$
$$\sigma_{xx}^{(1)}(L_1, t) = 0, \quad \sigma_{xx}^{(2)}(L_1 + g_0, t) = 0, \quad t \in t_{\text{sep}}, \quad (11)$$

while in the case of the mechanical contact we obtain

$$u_1(0, t) = 0, \quad u_2(L_1 + g_0 + L_2, t) = 0,$$
$$g(t) = g_0 + u_2(L_1 + g_0, t) - u_1(L_1, t) = 0,$$
$$\sigma_{xx}^{(1)}(L_1, t) = \sigma_{xx}^{(2)}(L_1 + g_0, t) = -p(t), \; t \in t_{\text{con}}, \quad (12)$$

where $\theta_n(x, t)$ is the temperature in the layers with respect to the temperature of the natural state, $t_{\text{sep}} \cup t_{\text{con}} \equiv (0, \infty)$, where $t_{\text{sep}}, t_{\text{con}}$ are, so far unknown, segments of time. From the solution of Eq. (7) with the boundary conditions (11) we would obtain the magnitude of the gap $g(t)$ while from the solution of Eq. (7) with the boundary conditions (12) the contact pressure $p(t)$. The system of partial differential Eqs. (7) and (8) is coupled, linear, elliptic-parabolic type. The boundary conditions are nonlinear, and coupled through the solution. The coupling of the partial differential equations of thermoelasticity is negligible for processes which are slow in time. Therefore we neglect the term with the time derivative in Eq. (8) getting

$$\frac{\partial^2\theta_n}{\partial x^2} - \frac{1}{k_n}\frac{\partial\theta_n}{\partial t} = 0,$$
$$\text{for} \quad n = 1, \quad x \in (0, L_1), \quad \text{for} \quad n = 2,$$
$$x \in (L_1 + g_0, L_1 + g_0 + L_2), \quad t \in (0, \infty). \quad (13)$$

We introduce the function

$$q(t) = g_0 - \tilde{\alpha}_1\int_0^{L_1}\theta_1(\xi, t)d\xi$$
$$- \tilde{\alpha}_2\int_0^{L_2}\theta_2(\xi + L_1 + g_0, t)d\xi \quad t \in (0, \infty). \quad (14)$$

We have

$$t \in t_{\text{sep}}, \quad q(t) \equiv g(t) > 0,$$
$$\text{while for} \quad t \in t_{\text{con}}, \quad q(t) \equiv -\frac{p(t)}{p_0} \quad (15)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$\tilde{\alpha}_n = \frac{1+\nu_n}{1-\nu_n}\alpha_n, \quad \tilde{E}_n = \frac{1-\nu_n}{(1+\nu_n)(1-2\nu_n)}E_n,$$

valid for $\epsilon_{yy}^{(n)} = \epsilon_{zz}^{(n)} = 0$,

while

$\tilde{\alpha}_n = \alpha_n, \tilde{E}_n = E_n$, valid for

$$\sigma_{yy}^{(n)} = \sigma_{zz}^{(n)} = 0, \quad p_0 = \left(\frac{L_1}{\tilde{E}_1} + \frac{L_2}{\tilde{E}_2}\right)^{-1}.$$

The thermal boundary conditions assume the form

$$\theta_1(0, t) = \vartheta_1\phi_1(t), \quad \text{for} \quad t \in (0, \infty),$$

$$K_1\frac{\partial\theta_1}{\partial x} = K_2\frac{\partial\theta_2}{\partial x}, \quad x = L_1, \quad t \in (0, \infty), \tag{16}$$

$$K_1\frac{\partial\theta_1}{\partial x} = \frac{\theta_2 - \theta_1}{R(q)}, \quad x = L_1, \quad t \in (0, \infty),$$

$$\theta_2(L_1 + g_0 + L_2, t) = \vartheta_2\phi_2(t), \quad t \in (0, \infty), \tag{17}$$

and the initial conditions

$$\theta_n(x, 0) = 0, \quad n = 1, 2. \tag{18}$$

$q(t)$ is a function of time and constitutes the integral characteristics of the problem considered. Now we are in position to write

$$R(q) = R_{\text{sep}}(q)H(q) + R_{\text{con}}(-p_0q)H(-q). \tag{19}$$

Thus it is evident that the thermal resistance taken into account in the boundary conditions of the problems of thermoelasticity leads to quantitatively and qualitatively new problems with essentially nonlinear boundary conditions.

3 Solution of the Nonstationary Problem

The geometry of the problem we solve is presented in Fig. 1. In order to solve the partial differential Eq. (13) with boundary and initial conditions (16)–(18) we can reduce it to an equivalent integral equation, apply Biot's method, asymptotic solutions of the problems with nonlinear boundary conditions or the method of finite difference. Here we apply the method of consecutive approximations. We assume as the jth approximation the analytical solution of the following boundary value problem:

$$k_n\frac{\partial^2\theta_n^{(j)}(x, t)}{\partial x^2} = \frac{\partial\theta_n^{(j)}(x, t)}{\partial t}, \quad n = 1, 2, \quad j = 1, 2, \ldots,$$

$$\theta_1^{(j)}(0, t) = \vartheta_1\phi_1(t), \quad t \in (0, \infty)$$

$$K_1\frac{\partial\theta_1^{(j)}(x, t)}{\partial x} = K_2\frac{\partial\theta_2^{(j)}(x, t)}{\partial x}, \quad x \in L_1, \quad t \in (0, \infty),$$

$$K_1\frac{\partial\theta_1^{(j)}(x, t)}{\partial x} = \frac{\theta_2^{(j)}(x, t) - \theta_1^{(j)}(x, t)}{R_i^{(j-1)}},$$

$$x \in L_1, \quad t \in (0, \infty),$$

$$\theta_2^{(j)}(L_1 + g_0 + L_2, t) = \vartheta_2\phi_2(t), \quad t \in (0, \infty),$$

$$\theta_n^{(j)}(x, 0) = 0. \tag{20}$$

We begin the process of computation from the calculation of all quantities at the moment of time t_i.

The First Iteration: Thermal resistance is $R_1^{(0)} = R(g_0) =$ const, next we find $\theta_1^{(1)}(x, t_1)$ and $\theta_2^{(1)}(x, t_1)$ using formulae (28) and $q_1^{(1)}(t_1)$ by Eq. (29).

The Second Iteration: We compute the thermal resistance $R_1^{(1)} = R(q_1^{(1)})$, then we find $\theta_1^{(2)}(x, t_1)$ and $\theta_2^{(2)}(x, t_1)$ using formulae (28), and $q^{(2)}(t_1)$ by Eq. (29).

The j_1th Iteration: We compute the thermal resistance $R_1^{(j_1-1)} = R(q_1^{(j_1-1)})$, next we find $\theta_1^{(j_1)}(x, t_1)$ and $\theta_2^{(j_1)}(x, t_1)$ using formulae (28), and $q^{(j_1)}(t_1)$ by Eq. (29). The process is terminated when the conditions for the j_1th iteration have been fulfilled.

Next we pass to the moment of time t_2. Similarly as for t_1 we find in turn:

The First Iteration: Thermal resistance is $R_2^{(0)} = R(q^{(j_1)}(t_1))$, next we find $\theta_1^{(1)}(x, t_2)$ and $\theta_2^{(1)}(x, t_2)$ using formulae (28) and $q_1^{(1)}(t_2)$ by Eq. (29).

The Second Iteration: We compute the thermal resistance $R_2^{(1)} = R(q_2^{(1)})$, then we find $\theta_1^{(2)}(x, t_2)$ and $\theta_2^{(2)}(x, t_2)$ using formulae (28), and $q^{(2)}(t_2)$ by Eq. (29).

The j_2th Iteration: We compute the thermal resistance $R_2^{(j_2-1)} = R(q_2^{(j_2-1)})$, next we find $\theta_1^{(j_2)}(x, t_2)$ and $\theta_2^{(j_2)}(x, t_2)$ using formulae (28), and $q^{(j_2)}(t_2)$ by Eq. (29). The process is completed when the conditions for the j_2th iteration have been fulfilled.

For the moment of time t_i. Similarly as above we find in turn:

The First Iteration: Thermal resistance is $R_i^{(0)} = R(q^{(j_{i-1})}(t_{i-1}))$, next we find $\theta_1^{(1)}(x, t_i)$ and $\theta_2^{(1)}(x, t_i)$, and $q_1^{(1)}(t_i)$.

The Second Iteration: We compute the thermal resistance $R_i^{(1)} = R(q_i^{(1)})$, then we find $\theta_1^{(2)}(x, t_i)$ and $\theta_2^{(2)}(x, t_i)$, and $q^{(2)}(t_i)$,

The j_ith Iteration: We compute the thermal resistance $R_i^{(j_i-1)} = R(q_i^{(j_i-1)})$, next we find $\theta_1^{(j_i)}(x, t_i)$ and $\theta_2^{(j_i)}(x, t_i)$, and $q^{(j_i)}(t_i)$. The process is terminated when the conditions for the j_ith iteration have been fulfilled.

The first approximation is computed at the moment of time t_i, $i = 1, 2, \ldots$ for $R_i^{(0)} =$ const equal to the magnitude of the thermal resistance at the preceding moment of time t_{i-1}, $R_i^{(0)} = R_{i-1}^{(j_{i-1})}$, $i = 2, 3, \ldots$, $R_1^{(0)} = R(g_0)$. The consecutive approximations for $j = 2, 3, \ldots, j_i$ are taken according to the rule

$$R_i^{(j-1)} = R(q_i^{(j-1)}), \quad q_i^{(j)} = q^{(j)}(t_i) = \text{const},$$

$$q^{(j)}(t) = g_0 - \tilde{\alpha}_1\int_0^{L_1}\theta_1^{(j)}(\xi, t)d\xi - \tilde{\alpha}_2\int_0^{L_2}\theta_2^{(j)}(\xi + L_1 + g_0, t)d\xi, \tag{21}$$

where t_i denotes the time for which we have to find the change of temperature in the considered thermal system. We introduce a net with respect to time, the nodal points of the net correspond to points t_i, $i = 1, 2, \ldots, t_0 = 0$, the time step is $t_i - t_{i-1}$. For time t_i we complete the iterational process on the j_ith iteration when all the following conditions are simultaneously satisfied:

$$\left|\frac{q_i^{(j_i)} - q_i^{(j_i-1)}}{q_i^{(j_i)}}\right| \cdot 100\% < \epsilon_q, \quad \left|\frac{\theta_n^{(j_i)} - \theta_n^{(j_i-1)}}{\theta_n^{(j_i)}}\right| \cdot 100\% < \epsilon_{\theta_n},$$

$$\left|\frac{R_i^{(j_i)} - R_i^{(j_i-1)}}{R_i^{(j_i)}}\right| \cdot 100\% < \epsilon_R.$$

To solve the problem we apply to the set of Eqs. (20) and (21) the Laplace transforms with respect to time variable. We introduce the following symbols:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$(\bar{\theta}_n^{(j)}(x, s), \bar{q}^{(j)}(s)) = \mathfrak{L}[\theta_n^{(j)}(x, t), q^{(j)}(t); t \to s] \equiv \int_0^\infty (\theta_n^{(j)}(x, t), q^{(j)}(t)) \exp(st)dt. \quad (22)$$

The solutions in the space of Laplace transforms assume the form

$$\Delta^{(j-1)}(s)\bar{\theta}_n^{(j)}(x, s) = \vartheta_1\bar{\phi}_1(s)T_{n1}(x, s) + \vartheta_2\bar{\phi}_2(s)T_{n2}(x, s), \quad n = 1, 2,$$

$$\Delta^{(j-1)}(s)\bar{q}^{(j)}(s) = g_0S^{-1}\Delta^{(j-1)}(s) - \vartheta_1\bar{\phi}_1(s)Q_1(s) - \vartheta_2\bar{\phi}_2(s)Q_2(s) \quad (23)$$

where

$$T_{11}(x, s) = R_i^{(j-1)} \cosh q_1(L_1 - x) \cosh q_2L_2 + \frac{R_1}{q_1L_1} \sinh q_1(L_1 - x) \cosh q_2L_2 + \frac{R_2}{q_2L_2} \cosh q_1(L_1 - x) \sinh q_2L_2,$$

$$T_{12}(x, s) = \frac{R_1}{q_1L_1} \sinh q_1x, \; x \in (0, L_1),$$

$$T_{21}(x, s) = \frac{R_2}{q_2L_2} \sinh q_2(L_1 + L_2 + g_0 - x), \quad (24)$$

$$T_{22}(x, s) = R_i^{(j-1)} \cosh q_1L_1 \cosh q_2(x - L_1 - g_0) + \frac{R_1}{q_1L_1} \sinh q_1L_1 \cosh q_2(x - L_1 - g_0) + \frac{R_2}{q_2L_2} \cosh q_1L_1 \sinh q_2(x - L_1 - g_0), \quad x \in (L_1 + g_0, L_1 + g_0 + L_2),$$

$$Q_n(s) = \tilde{\alpha}_1L_1Q_{n1}(s) + \tilde{\alpha}_2L_2Q_{n2}(s),$$

$$q_1L_1Q_{11}(s) = R_i^{(j-1)} \sinh q_1L_1 \cosh q_2L_2 + \frac{R_1}{q_1L_1} (\cosh q_1L_1 - 1) \cosh q_2L_2 + \frac{R_2}{q_2L_2} \sinh q_1L_1 \sinh q_2L_2,$$

$$(q_2L_2)^2Q_{12}(s) = R_2(\cosh q_2L_2 - 1), \; (q_1L_1)^2Q_{21}(s) = R_1(\cosh q_1L_1 - 1),$$

$$q_2L_2Q_{22} = R_i^{(j-1)} \sinh q_2L_2 \cosh q_1L_1 + \frac{R_1}{q_1L_1} \sinh q_1L_1 \sinh q_2L_2 + \frac{R_2}{q_2L_2} (\cosh q_2L_2 - 1) \cosh q_1L_1,$$

$$\Delta^{(j)}(s) = R_i^{(j)} \cosh q_1L_1 \cosh q_2L_2 + \frac{R_1}{q_1L_1} \sinh q_1L_1 \cosh q_2L_2 + \frac{R_2}{q_2L_2} \cosh q_1L_1 \sinh q_2L_2,$$

$$q_n = \sqrt{\frac{s}{k_n}}, \quad R_n = \frac{L_n}{K_n}, \quad n = 1, 2. \quad (25)$$

Let us consider the state of the system of thermal contact in the case when the temperature of the walls varies according to the law

$$\phi_n(t) = [1 + \psi_n \exp(-h^2t)]H(t), \text{ i.e.,}$$

$$\bar{\phi}_n(s) = \frac{s(1 + \psi_n) + h^2}{s(s + h^2)}. \quad (26)$$

For $t \to 0$, $\phi_n(t) \to 1 + \psi_n$, at the initial time the temperature of the walls is therefore equal to $(1 + \psi_n)\vartheta_n$. For $t \to \infty$, $\phi_n(t) \to 1$, i.e., when the temperature of walls is constant, ϑ_n. Such an assumption of the walls temperatures makes it possible to investigate the time behavior of the contact characteristics (i.e., the gap, contact pressure, contact temperatures, and thermal resistance) since the initial state, for a small gap, to the exit on the steady-state. The steady-states and the behavior of the processes resulting from small perturbations in the vicinities of such states were investigated by Barber and Zhang (1988). We have shown that steady-state solution can be unique and nonunique, as well as stable and nonstable. On the other hand the assumed choice of the thermal loading furnishes the possibility of simple control of the process and the exit of the system on different steady-states in the case of the uniqueness of the stationary solution, the uniqueness and stability of the stationary solution, and the uniqueness and nonstability of the stationary solution.

The inverse Laplace transforms of the jth approximation can be calculated from the corresponding formulae (see, for example, Carslaw and Jaeger, 1959). The required fields will be presented in the form of the series each term of which corresponds to the residuum of the Laplace transform in terms of the characteristic equation:

$$\Delta^{(j)}(s_k) = 0, \quad s_k = -x_k^2 \frac{k_1}{L_1^2}, \quad \Im s_k = 0, \quad \Re s_k < 0, \quad (27)$$

with respect to the poles $s = 0$ and $s = -h^2$ of functions $\bar{\phi}_n(s)$, given by Eqs. (26).

The time behavior of the surface in contact $\theta_n^0(t)$, $(\theta_n^{(j)}(L_1, t) \to \theta_n^0(t), n = 1, 2)$ and that of function $q(t)$, $(q^{(j)}(t) \to q(t))$ has a physical meaning (see Eq. (15)) and can be written as follows:

$$\theta_n^{(j)}(L_1, t) = \hat{\theta}_n^{(j)} + \sum_{k=0}^{\infty} \frac{T_{nk}^{(j)}}{Z_k^{(j)}} \exp(-x_k^2F_1^0), \quad j = 1, 2, 3, \ldots, \quad (28)$$

$$q^{(j)}(t) = \hat{Q}^{(j)} - \tilde{\alpha}_2L_2 \sum_{k=0}^{\infty} \frac{Q_k^{(j)}}{Z_k^{(j)}} \exp(-x_k^2F_1^0), \quad j = 1, 2, \ldots, \quad (29)$$

where

$$\hat{\theta}_n^{(j)} = [\vartheta_n(R_i^{(j-1)} + R_p) + \vartheta_pR_n]f_i^{(j-1)}, \quad n, p = 1, 2, \quad n \neq p, \quad S = R_1 + R_2, \quad (30)$$

$$\hat{Q}^{(j)} = Q_\infty + \beta f_i^{(j-1)}, \quad (31)$$

$$Q_\infty = g_0 - u, \; u = \tilde{\alpha}_1L_1\vartheta_1 + \tilde{\alpha}_2L_2\vartheta_2, \; 2\beta = \alpha L(\vartheta_1 - \vartheta_2), \; \alpha L = \tilde{\alpha}_1L_1R_1 - \tilde{\alpha}_2L_2R_2, \; f_i^{(j-1)} = f(q_i^{(j-1)}),$$

$$f(q) = \frac{1}{S + R(q)}, \quad (32)$$

$$T_{nk}^{(j)} = [\theta_{nk}(R_i^{(j-1)}C_p + R_pS_p) + \theta_{pk}R_nS_n]|_{x=x_k},$$

$$Q_k^{(j)} = \{\theta_{1k}[\alpha(R_i^{(j-1)}S_1C_2 + R_1\tilde{C}_1C_2 + R_2S_1S_2) + R_2\tilde{C}_2] + \theta_{2k}[R_i^{(j-1)}S_2C_1 + R_2\tilde{C}_2C_1 + R_1S_1S_2 + \alpha R_1\tilde{C}_1]\}|_{x=x_k},$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\theta_{nk} = \frac{\vartheta_n}{1 + \psi_n(1 - n_1^2/x_k^2)}, \quad n, p = 1, 2,$$

$$n \neq p, \quad k = 1, 2, \ldots,$$

$$\theta_{n0} = \vartheta_n \psi_n, \; Z_k^{(j)} = \frac{1}{2} x_k \left[\frac{\partial}{\partial x} Z(x) \right] \Big|_{x=x_k},$$

$$k = 1, 2, \ldots, \quad (33)$$

$$Z_0^{(j)} = Z(x_0), \; Z(x) = R_i^{(j-1)} C_1 C_2 + R_1 S_1 C_2 + R_2 C_1 S_2,$$

$$C_1 = \cos x, \quad C_2 = \cos \tilde{n} x, \quad S_1 = \frac{\sin x}{x}, \quad S_2 = \frac{\sin \tilde{n} x}{\tilde{n} x}$$

$$\tilde{C}_1 = \frac{\cos x - 1}{x^2}, \quad \tilde{C}_2 = \frac{\cos \tilde{n} x - 1}{(\tilde{n} x)^2},$$

$$\alpha = \frac{\tilde{\alpha}_1 L_1}{\tilde{\alpha}_2 L_2}, \quad F^0 L_1^2 = k_1 t,$$

$$n_p = h \frac{L_p}{\sqrt{k_p}}, \quad \tilde{n} = \frac{n_2}{n_1}, \quad Z(x_k) = 0, \; k = 1, 2, \ldots, \quad x_0 = n_1.$$

Let us note that the first terms $\hat{\theta}_n^{(j)}$, $\hat{Q}^{(j)}$ in Eqs. (28), (29) are the residua of the corresponding Laplace's transforms with respect to the pole $s = 0$ and they refer to the temperatures of the planes in contact for the stationary problem, i.e., for $t \to \infty$.

4 On the Convergence of the Iterative Process

The iterative scheme (21) and (29) can be written in the following form:

$$q_i^{(j)} = Q(q_i^{(j-1)}, t_i), \quad j = 2, 3, \ldots, \quad (34)$$

and for

$$F_i^0 \equiv \frac{k_1 t_i}{L_1^2} \gg 1, \quad \hat{q}^{(j)} = \hat{Q}(\hat{q}^{(j-1)}),$$

$$\hat{Q}(q) = Q_\infty + \beta f(q). \quad (35)$$

Function $Q(q, t_i)$ depends also on all thermomechanical characteristics of the model, this fact makes the analytical proof of the iterative process difficult and complicated. We shall show the existence of the solutions and shall verify the convergence of the process by computational simulation.

Let the numerical scheme $q_i^{(j)} \to q_i$ (i.e., $\hat{q}^{(j)} \to \hat{q}$) converge to the root q_i, ($\hat{q}$) of the equation

$$q_i = Q(q_i, t_i), \quad (\hat{q} = \hat{Q}(\hat{q})), \quad (36)$$

provided the following conditions are satisfied:

$$\left| \frac{d}{dq} Q(q_i, t_i) \right| < 1, \quad \left(\left| \frac{d}{dq} \hat{Q}(\hat{q}) \right| < 1 \right). \quad (37)$$

Point $q_i(\hat{q})$ will be called the *fixed point* of the function (precisely its mapping) $Q \to \hat{Q}$ if condition (36) is satisfied, and *fixed stable point* provided both conditions (36) and (37) are fulfilled.

The computational simulation shows us that for certain values of the mechanical and physical parameters, and for $\beta > 0$ there exists for mapping (36), up to a certain moment of time $t_i < t_*$, the fixed, stable point q_i. Beginning from a certain time t_* we obtain a bifurcation, the fixed point q_i looses its stability and at its neighborhood two new stable fixed points appear, say q_i^1 and q_i^2:

$$q_i^{(2j-1)} \to q_i^1, \quad q_i^{(2j)} \to q_i^2, \quad t_i > t_*,$$

$$q_i^1 = Q(q_i^2, t_i), \quad q_i^2 = Q(q_i^1, t_i). \quad (38)$$

We note that the reverse situation is also possible. We can obtain

$$\text{for} \quad F_i^0 \equiv \frac{k_1 t_i}{L_1^2} \gg 1, \quad q_i^1 \to \hat{q}^1, \quad q_i^2 \to \hat{q}^2, \quad (39)$$

and

$$\hat{q}^1 = \hat{Q}(\hat{q}^2), \quad \hat{q}^2 = \hat{Q}(\hat{q}^1),$$

$$\hat{q}^1 < \hat{q} < q^2, \quad \hat{q} = \hat{Q}(\hat{q}), \quad \frac{d}{dq} \hat{Q}(\hat{q}) < -1. \quad (40)$$

Since the two convergent subsequences (4) $q_i^{(2j-1)} \to q_i^1$, $q_i^{(2j)} \to q_i^2$ tend to two different points, therefore the iterational sequence $q_i^{(j)}$ becomes divergent.

5 Function of Thermal Resistance

The determination of the thermal surface resistance of two solids in contact was the subject of the experimental and theoretical investigations (Barber and Zhang, 1988; Cooper et al., 1969; Clausing and Chao, 1965; Duvaut, 1979; Shlykov and Ganin, 1964; Srinivasan and France, 1985). We can approximate the results by means of the function of thermal resistance in the form

$$R(q) = \begin{cases} R(0) r \dfrac{q}{R(0) K_{\text{air}}}, & q > 0, \\ R(0) r^{-1}(-\alpha_* p_0 q), & q < 0, \end{cases} \quad (41)$$

where

$$r(x) = 1 + x.$$

$R(0)$ denotes the thermal resistance corresponding to the zero contact pressure, α_* is a parameter of sensitivity of thermal resistance to the contact pressure, and K_{air} is the heat conduction of air in the gaps between the surfaces. The function of thermal resistance fulfills the following requirements: it has to be a good approximation to the experimental data for small number of parameters. The proposed algorithm of solution holds also in the case of a pointwise interpolation of the experimental results.

6 Conclusions

The algorithm considered above can be applied not only to stationary and nonstationary problems but also to dynamical problems with the inertia terms in (4) taken into account. Then the system of Eqs. (4) and (5) is of hyperbolic-parabolic type and the procedure of numerical computation becomes very tedious. It was shown that the iterative process is convergent and we have found the conditions (inequalities) which guarantee the convergence. We also have shown when a bifurcation in the solution and the conditions of stability may appear. We haven't discussed the very special cases of the thermal resistance function for which the solution of the thermoelasticity problem is unique (compare Duvaut, 1979).

The statement about the existence of the stationary solutions results from the construction of the stationary solutions. The statements about stability and nonstability of steady-state solutions have been discussed by Barber and Zhang (1988). The conditions of nonstability of intermediate steady-state and nonunique solution coincide with the conditions (43) of the nonstability of the fixed point of the mapping. The condition of the nonstability of the unique steady-state solution (44) may serve as a condition for choosing the model parameters such that the iteration sequence is divergent while the subsequences converge. In the theory of nonlinear equations, for example in hydromechanics such conditions are called the conditions of the

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

turbulence initiation. In the dynamical systems they correspond to the conditions of the initiation of chaotic motion.

Acknowledgments

The postdoctoral scholarship of the Ministry of Education of Poland for one of the authors (Yu. A. P.) is gratefully acknowledged. This support enabled him to work during the academic year 1992–1993 at the University of Warsaw.

The authors are thankful to the reviewer of the paper for his exhaustive comments and suggestions.

References

Barber, J. R., 1973, "Indentation of the Semi-infinite Elastic Solid by a Hot Sphere," *International Journal of Mechanical Science,* Vol. 15, pp. 813–819.

Barber, J. R., 1978, "Contact Problems Involving a Cooled Punch," *Journal of Elasticity,* Vol. 8, pp. 409–423.

Barber, J. R., and Comninou, Maria, 1989, "Thermoelastic Contact Problems," *Thermal Stresses,* R. B. Hetnarski, ed., Elsevier Science Publ. B. V., Amsterdam.

Barber, J. R., Dundurs, J., and Comninou, Maria, 1979, "Stability Considerations in Thermoelastic Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 871–874.

Barber, J. R., and Zhang, R., 1988, "Transient Behaviour and Stability for the Thermoelastic Contact of Two Rods of Dissimilar Materials," *International Journal of Mechanical Science,* Vol. 30, pp. 691–704.

Boussinesq, J. V., 1885, *Applications des Potentiels à l'Étude de l'Équilibre et du Mouvement des Solides Élastiques,* Paris.

Carslaw, H. S., and Jaeger, J. C., 1959, *The Conduction of Heat in Solids,* 2nd ed., Clarendon Press, Oxford, U.K.

Clausing, A. M., and Chao, B. T., 1965, "Thermal Contact Resistance in a Vacuum Environment," ASME *Journal of Heat Transfer,* Vol. 87, pp. 243–251.

Comninou, Maria, and Dundurs, J., 1979, "On the Barber Boundary Conditions for Thermoelastic Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 849–853.

Cooper, A. M., Mikić, B. B., and Yovanovich, M. M., 1969, "Thermal Contact Conductance," *Int. J. Heat Mass Transfer,* Vol. 12, p. 279.

Dundurs, J., and Comninou, Maria, 1976, "On the Boundary Conditions in Contact Problems with Heat Conduction, *Developments in Theoretical and Applied Mechanics,* Vol. 8, R. P. McNitt, ed., Blacksburg, VA.

Duvaut, G., 1979, "Free Boundary Problem Connected with Thermoelasticity and Unilateral Contact," *Free Boundary Problems,* Vol. 2, F. Severi, ed., Inst. Naz. Alta Mat., Pavia, pp. 217–236.

Gladwell, G. M. L., 1980, *Contact Problems in the Classical Theory of Elasticity,* Sijthoff and Noordhoff, Alphen aan den Rijn.

George, D. L., and Sneddon, I. N., 1962, "The Axisymmetric Boussinesq Problem for a Heated Punch," *Journal of Mathematics and Mechanics,* Vol. 11, pp. 665–689.

Goryacheva, I. G., and Dobychin, M. N., 1988, *Contact Problems in Tribology,* Mashinostroenye, Moscow (in Russian).

Hertz, H., 1881, "Über die Berührung fester elastischer Körper," *Journal für reine und angewandte Mathematik,* Vol. 92, pp. 156–171; also in *Gesammelte Werke,* in 3 volumes, Leipzig, 1894.

Johnson, K. L., 1987, *Contact Mechanics,* Cambridge University Press, Russian ed., Mir, Moscow, 1989.

Nowacki, W., 1986, *Thermoelasticity,* 2nd ed., PWN—Pergamon Press, Warsaw.

Olesiak, Z. S., 1993, "O zagadnieniach, w których naprężenia w istotny sposób zależą od kierunku strumienia ciepla," *Zagadnienia Maszyn Przeplywowych,* Gdańsk, IMP Polish Academy of Science (in Polish).

Olesiak, Z., 1981, "On a Method of Solution of Mixed Boundary-value Problems of Thermoelasticity," *Journal of Thermal Stresses,* Vol. 4, pp. 501–508.

Shlykov, Yu. P., and Ganin, Ye, A., 1964, "Thermal Resistance of Metallic Contacts," *International Journal of Heat Mass Transfer,* Vol. 7, pp. 921–929.

Srinivasan, M. G., and France, D. M., 1985, "Non-uniqueness in Steady-state Heat Transfer in Prestressed Duplex Tubes—Analysis and case History," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 257–262.

Thomas, T. R., and Probert, S. D., 1970, "Thermal contact resistance. The directional effect and other problems," *International Journal of Heat Mass Transfer,* Vol. 13, pp. 789–807.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

On Nonuniqueness and Stability in Barber's Model of Thermoelastic Contact

Z. S. Olesiak

Yu. A. Pyryev[1]

Department of Applied Mathematics and Mechanics,
University of Warsaw
Warsaw, Poland

In the recent our paper we have derived an algorithm which lets us find numerical solutions of special cases of a nonstationary problem of thermoelasticity with imperfect boundary conditions of Barber's model. In this paper we show the results of computational simulation for a case of the thermal resistance function. We have obtained a number of solutions for different situations and have discussed the unique, nonunique, stable, and unstable solutions. We have found cases when the unique and nonunique solutions alternate. The results have been presented in the form of diagrams.

1 Introduction

In this paper we apply the computational simulation to find numerical solutions of a number of special cases of a nonstationary problem of thermoelasticity with imperfect boundary conditions. We make use of an algorithm derived by Olesiak and Pyryev (1996). The problem is that of two layers of different materials with the initial gap between them and corresponds to two contacting rods investigated by Barber and Zhang (1988). The assumed choice of the thermal loading furnishes the possibility of simple control of the process and the exit of the system on different steady-states in the case of the uniqueness of the stationary solution (point 2.4), the uniqueness and stability of the stationary solution (point 2.5), and the uniqueness and nonstability of the stationary solution (point 2.6). The results have been presented in the diagram form. The discussion is summarized in the Conclusion.

2 Numerical Investigation of the Transient Process

We consider the behavior of the thermoelastic contact characteristics since the moment of the beginning of the thermal disturbance to the exit of the system on the steady-state process. We show, for a special case of the thermal resistance function (point 2.1), by varying the temperatures of the walls and the value of the initial gap that the nonstationary solution tends to one of the nonunique steady-state solutions (point 2.4). In the case when the characteristics of the contact tend to the unique steady-state solution (point 2.5) we investigate the influence of the intensity of variation of the thermal resistance on pressure. In point 2.6 we show the behavior of the solution in the case of the uniqueness and instability of the steady-state solution.

2.1 Function of Thermal Resistance. The determination of the thermal surface resistance of two solids in contact was the subject of the experimental and theoretical investigations (Barber and Zhang, 1988; Cooper, Mikić, and Yovanovich, 1969; Clausing and Chao, 1965; Duvaut, 1979; Shlykov and Ganin, 1964; Srinivasan and France, 1985; the references are given in Olesiak and Pyryev, 1996). The function of thermal resistance can be assumed in the form

$$R(q) = \begin{cases} R(0)r\left(\dfrac{q}{R(0)K_{\text{air}}}, a_+, b_+, c_+, d_+\right), & q > 0, \\ R(0)r^{-1}(-\alpha_* p_0 q, a_-, b_-, c_-, d_-), & q < 0, \end{cases} \tag{1}$$

where

$$r(x, a, b, c, d) = 1 + x\frac{1 + ax^c}{1 + bx^d}, \quad c + 1 > d,$$

$R(0)$ denotes the thermal resistance corresponding to the zero contact pressure, α_* is a parameter of sensitivity of thermal resistance to the contact pressure, and K_{air} is the heat conduction of air in the gaps between the surfaces. The parameters $a_\pm$, $b_\pm$, $c_\pm$, $d_\pm$, and the form of function $r(.)$ are such that we obtain good approximations of the thermal resistance for a bigger gap $g = q$, $(q > 0)$ and the contact pressure $p = -p_0 q$, $(q < 0)$. The function of thermal resistance fulfills the following requirements: it has to be a good approximation to the experimental data for small number of parameters. The proposed algorithm of solution holds also in the case of a pointwise interpolation of the experimental results. In our calculations we assume a simplified form of $r(x) = 1 + x$; i.e., we take $a_\pm = b_\pm = 0$. Such a form of the thermal resistance is sufficient for our purpose to show the characteristic features of the process. We do not discuss the dependance of the solutions on the character of approximation of thermal resistance functions, and on their smoothness.

2.2 Solutions for Stationary Problems. First we shall discuss the states of the system which lead to a stationary process. It was shown in Barber (1978), Barber and Zhang (1988), and Barber, Dundurs, and Comninou (1979) (see the references in Olesiak and Pyryev, 1996) that the stationary problem may exhibit both: the unique and nonunique solutions. In the case of special forms of the function of thermal resistance, Eq. (1), we can write down the conditions that should satisfy the initial parameters of the system such that the stationary problem has the unique solution. In the case of the conditions

$$\beta < 0, \quad g_0^- \le g_0 \le g_0^+ \tag{2}$$

where

$$g_0^- = \begin{cases} u + 2\sqrt{-\beta K_{\text{air}}} - \dfrac{K_{\text{air}}}{f(0)}, & \beta_+ > 0, \\ u - \beta f(0), & \beta_+ < 0, \end{cases}$$

[1] Permanent address: Chair of Mechanics, University of Lviv, Lviv, Ukraine.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Sept. 27, 1993; final revision, May 12, 1995. Associate Technical Editor: F. Y. M. Wan.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

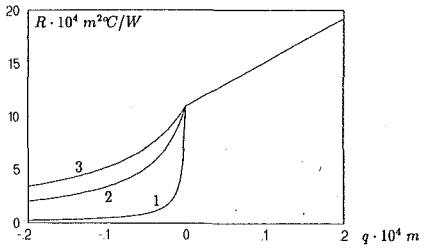


Fig. 1 Thermal resistance R as a function of $g \equiv q$ ($q > 0$) and of contact pressure $p \equiv -p_0 q$ ($q < 0$) for α_*: 1 – $1 \cdot 10^{-5}$, 2 – $1 \cdot 10^{-6}$, 3 – $5 \cdot 10^{-7}$ Pa^{-1}

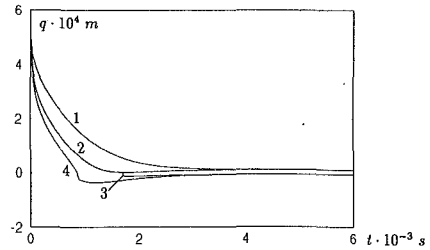


Fig. 3 Gap $g \equiv q$ ($q > 0$) and contact pressure $p \equiv -p_0 q$ ($q < 0$) versus time, for a set of values of thermal induction $\vartheta_n(1 + \psi_n)$; $\vartheta_1 = 300$°C, $\vartheta_2 = 0$°C, $\vartheta_1\psi_1$: 1 – 0°C, 2 – 221°C, 3 – 222°C, 4 – 320°C

$$g_0^{+} = \begin{cases} u + \left(\sqrt{-\frac{\beta}{S}} - \sqrt{\frac{R(0)}{S\alpha_* p_0}}\right)^2 + \frac{1}{\alpha_* p_0}, & \beta_- > 0, \\ u - \beta f(0), & \beta_- < 0, \end{cases}$$

$$\beta_{\pm} = \beta \frac{df_{\pm}(0)}{dq} - 1, \quad \frac{df_{\pm}(0)}{dq} = \lim_{q \to 0\pm} f'_{\pm}(q$$

$$= \begin{cases} -\frac{f^2(0)}{K_{\text{air}}} S, \\ -f^2(0)\alpha_* p_0 R(0) S. \end{cases}$$

The equation (see Eq. (36) in Olesiak and Pyryev, 1996) $q = \hat{Q}(q)$ has more than one solution, namely for $g_0 = g_0^{\pm}$ there exist precisely two solutions while for $g_0 \in (g_0^-, g_0^+)$ three solutions, namely $\tilde{q}^1, \tilde{q}^0, \tilde{q}^2 (\tilde{q}^1 < \tilde{q}^0 < \tilde{q}^2)$. The intermediate solution $\tilde{q}^0$ turns out to be unstable

$$\frac{d}{dq}\hat{Q}(\tilde{q}^0) > 1 \quad \text{or} \quad \beta \frac{d}{dq} f(\tilde{q}^0) > 1. \tag{3}$$

For $\beta > 0$ the solution of the stationary problem is always unique, i.e., the root $\hat{q}$ of Eq. (36) (in Olesiak and Pyryev) is the fixed point of the mapping $\hat{Q}(q)$. In the case when at this point the condition is satisfied such that

$$\frac{d}{dq}\hat{Q}(\hat{q}) < -1, \quad \text{or} \quad \beta \frac{d}{dq} f(\hat{q}) < -1, \tag{4}$$

then point $\hat{q}$ is unstable and in its neighborhood there exist two stable points of the mapping $\hat{Q}(\hat{Q}(\hat{q}))$ such that $\hat{q}^1, \hat{q}^2$ ($\hat{q}_1 < \hat{q} < \hat{q}_2$) constitute the solutions of equations

$$\hat{q}^1 = \hat{Q}(\hat{q}^2), \quad \hat{q}_2 = \hat{Q}(\hat{q}^1). \tag{5}$$

The preceding solutions are those for which there exist two subsequences with two different points of accumulation and consequently we obtain the divergent iterational sequence $q_i^{(j)}$.

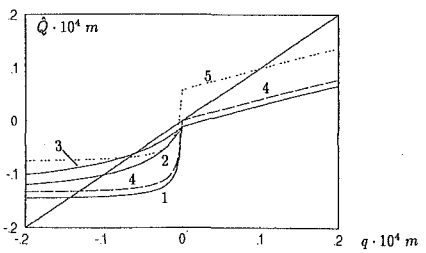


Fig. 2 Function $\hat{Q}$ versus q for the set of values g_0, α_*. The solution of Eq. (36) (in Olesiak and Pyryev) corresponds to the intersection of a curve and the straight line $\hat{Q} = q$; $g_0 = 4.99 \cdot 10^{-4}$ m, α_*: 1 – $1 \cdot 10^{-5}$, 2 – $1 \cdot 10^{-6}$, 3 – $5 \cdot 10^{-7}$ Pa^{-1}; and $\alpha_* = 1 \cdot 10^{-5}$, g_0: 4 – $5.001 \cdot 10^{-4}$, 5 – $5.06 \cdot 10^{-4}$ m.

2.3 Thermomechanical Parameters of the System of Thermal Contact. To discuss the problem analytically we make use of the parameters and data corresponding to those used by Barber and Zhang (1988). We take the values of thermal resistance: $R(0) = 0.11 \cdot 10^{-2}$ m²°C/W, $K_{\text{air}} = 2.43 \cdot 10^{-2}$ Wm^{-1}°C^{-1}. For these values the dependance of thermal resistance on parameter q has been shown in Fig. 1. Curves 1, 2, 3 correspond to the values of parameter $\alpha_* = 0.1 \cdot 10^{-4}$ Pa^{-1}, $0.1 \cdot 10^{-5}$ Pa^{-1}, and $0.5 \cdot 10^{-6}$Pa^{-1}. For the first layer we have assumed the following values of material constants: $\nu_1 = 0.3$, $\tilde{E}_1 = 0.19 \cdot 10^{12}$ Pa, $\tilde{\alpha}_1 = 0.14 \cdot 10^{-4}$°C^{-1}, $K_1 = 0.21 \cdot 10^2$ Wm^{-1}°C^{-1}, $k_1 = 0.59 \cdot 10^{-5}$ m² s^{-1}, and width $L_1 = 0.1$ m, while for the second layer. $\nu_2 = 0.32$, $\tilde{E}_2 = 0.72 \cdot 10^{11}$ Pa, $\tilde{\alpha}_2 = 0.22 \cdot 10^{-4}$°C^{-1}, $K_2 = 0.173 \cdot 10^3$ Wm^{-1}°C^{-1}, $k_2 = 0.672 \cdot 10^{-4}$ m² s^{-1}, and width $L_2 = 0.29$ m.

2.4 Transient Response in the Case of the Nonunique Solution of Stationary Problems. The numerical analysis of nonstationary behavior of thermal contact characteristics has been performed for thermal excitation of the system Eqs. (16), (17), and (36) (in Olesiak and Pyryev, 1996), for the values. $\vartheta_1 = 300$°C, $\vartheta_2 = 0$°C, $h^2 = 0.25 \cdot 10^{-2}$ s^{-1} and for the initial gap $g_0 = 0.506 \cdot 10^{-3}$ m. The law of the thermal resistance is represented by curve 1 in Fig. 1. Parameters of the system corresponding to the conditions (2) ($g_0^- = 0.5001 \cdot 10^{-3}$ m, $g_0^+ = 0.5090 \cdot 10^{-3}$ m) are such that the stationary solution is not unique. The graphical solution of equation $q = \hat{Q}(q)$ is shown in Fig. 2 by curve 5. The roots of the graphical solution are $\tilde{q}^1 = -0.686 \cdot 10^{-5}$ m, $\tilde{q}^2 = 1.000 \cdot 10^{-5}$, $\tilde{q}^0 = -4.54 \cdot 10^{-7}$.

The output of the system on various steady-states depends on the character of excitation of the walls temperatures. For the fixed value of h there exists the temperature of the initial excitation such that for higher values of the temperature the system alternates from one steady-state to another one. Thus for $\vartheta_1\psi_1 \leq 221$°C the system tends to the stationary solution with the gap $\tilde{q}^2$. For $\vartheta_1\psi_1 = 0$°C and $\vartheta_1\psi_1 = 221$°C the time behavior of the gap has been shown in Fig. 3 by curve 1 and curve 2, respectively. The time variation of the thermal resistance has been presented in Fig. 4 by curves 1 and 2, respectively. The time behavior of the temperature of the thermally contacting plane of the first layer $\theta_1^0(t)$ (continuous curves) and the temperatures of the second layer $\theta_2^0(t)$ (dashed curves) have been

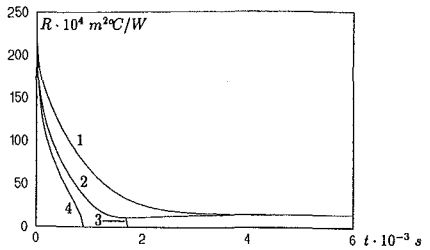


Fig. 4 Thermal resistance R versus time for the set of values of thermal induction $\vartheta_n(1 + \psi_n)$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

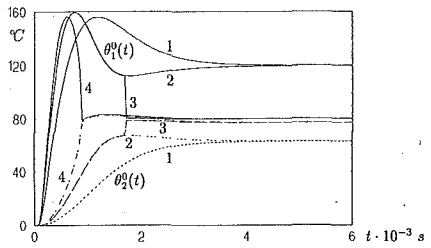


Fig. 5 Temperature of thermally contacting planes versus time for various values of thermal excitation $\vartheta_n(1 + \psi_n)$; $\vartheta_1 = 300°C$, $\vartheta_2 = 0°C$, $\vartheta_1\psi_1$: 1 – 0°C, 2 – 221°C, 3 – 222°C, 4 – 320°C.

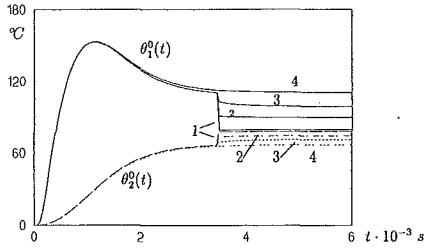


Fig. 7 Temperature of thermally contacting planes versus time for various values of g_0 and α_*, $g_0 = 4.99 \cdot 10^{-4}$ m, α_*: 1 – $1 \cdot 10^{-5}$, 2 – $1 \cdot 10^{-6}$, 3 – $5 \cdot 10^{-7}$ Pa^{-1}; 4 – $\alpha_* = 1 \cdot 10^{-5}$ Pa^{-1}, and $g_0 = 5.001 \cdot 10^{-4}$ m

shown in Fig. 5. The curves marked by "1" correspond to $\vartheta_1\psi_1 = 0°C$ while curves marked by "2" to $\vartheta_1\psi_1 = 221°C$, respectively.

For $\vartheta_1\psi_1 \geq 222°C$ the system tends to the stationary solution with $\tilde{q}^1$ (i.e., for contact pressure $p = -p_0 q^1$, $p_0 = 2.196 \cdot 10^{11}$ m^{-1} Pa).

On the other hand for $\vartheta_1\psi_1 = 222°C$ and for $\vartheta_1\psi_1 = 320°C$ the character of time variation of the gap $q > 0$ for $t \leq 1680$ s ($t \leq 840$ s) and the $q < 0$ for $t \geq 1720$ s ($t \geq 880$ s) have been shown in Fig. 3—curve 3, and curve 4, respectively.

The corresponding behavior of the thermal resistance for the same temperatures of both the layers is presented in Fig. 4 by curves 3 and 4, respectively. The character of the temperature behavior of the bounding plane in contact for the first layer (the continuous curve) and the second layer (dashed curve) in time have been shown, for the same temperatures, by curves 3 and 4, respectively, in Fig. 5.

Thus it is evident that the system in thermal contact enters, at a definite moment of time, in the mechanical contact. After that the characteristics of the thermoelastic contact are subjected to sharp changes, and finally, with time the stationary regime is reached.

The common temperature of the bounding planes still increases after entering into contact (see the curves 4 in Fig. 5). The time when the maximum temperature of the system is in contact exceeds therefore the time of the system without the mechanical contact.

2.5 Transient Response in the Case of the Unique Solution of Stationary and Nonstationary Problems. As an example we envisage the thermal excitation of the system as at point 2.4 for $\vartheta_n\psi_n = 0$. The value of the gap $g_0 = 0.499 \cdot 10^{-3}$ m corresponds to the condition of the existence of the unique stationary solution for $\beta < 0$, $g_0 \notin [g_0^-, g_0^+]$. The graphical solution of equation $q = \hat{Q}(q)$ from which we find the value of the stationary contact pressure is shown in Fig. 2 as the point of the intersection of curve 1 and the straight line $\hat{Q} = q$, $\hat{q} = -0.143 \cdot 10^{-4}$ m, $p = -p_0 q$. The time behavior of the gap $g = q(t \leq 3440$ s) and of tension q (≥ 3480 s) is presented in Fig. 6 by curve 1. The temperature of the planes in contact $\theta_n^0(t)$ is shown in Fig. 7 (curve 1). At the moment of time when the system comes into mechanical contact the thermal resistance changes sharply, and consequently the characteristics of the thermal contact change rapidly. In order to investigate the influence of the thermal resistance, more precisely the rate intensity of the thermal resistance on the contact pressure, we have performed calculations for the values $\alpha_* = 1 \cdot 10^{-6}$ Pa^{-1} and $\alpha_* = 5 \cdot 10^{-7}$ Pa^{-1}, the preceding parameters corresponding to the values $g_0^+ = 5.021 \cdot 10^{-4}$, $g_0^+ = 5.003 \cdot 10^{-4}$ m, respectively. In Fig. 7 the thermal resistances of the form 2 and 3 in Fig. 1 correspond to these values of α_*. The results of the computations are presented in Figs. 3, 7, 8 by curves 2 and 3, respectively.

At point 2.2 we pointed out that for $\beta < 0$, in the case of the unique solution, i.e., for $g_0 \notin [g_0^-, g_0^+]$, and the three nonunique solutions for $g_0 \in (g_0^-, g_0^+)$ of the stationary problems there may exist two nonunique solutions ($g_0 = g_0^+$).

The graphical solution of equation $q = \hat{Q}(q)$ is shown in Fig. 2 by the points of the intersection of the straight line $\hat{Q} = q$ and the curves 1–3, the straight line with the point curve 5, and the straight line with the dashed curve 4, respectively.

In order to illustrate the exit of the nonstationary process on one of the two stationary solutions, for example for $q = \tilde{q}^2 + 0(\tilde{q}^0 = \tilde{q}^2)$ the computations have been performed for the thermal excitation $\vartheta_1 = 300°C$, $\vartheta_2 = 0°C$, $\vartheta_n\psi_n = 0$ and $\alpha_* = 1 \cdot 10^{-5}$, $g_0 = g_0^- = 5.001 \cdot 10^{-4}$ m. The results of the computer simulations are shown by curves marked "4", in Figs. 6 and 7.

2.6 Transient Response in the Case of the Uniqueness and Nonstability of the Stationary Problems. The characteristics of the thermoelastic contact have been investigated for the initial gap $g_0 = 2.127 \cdot 10^{-3}$ m and $\alpha_* = 1 \cdot 10^{-5}$Pa^{-1}. The thermal perturbation of the system has been determined by the magnitudes $\vartheta_1 = 0°C$, $\vartheta_2 = 350°C$, and $\vartheta_2\psi_2 = 0°C$. For these entry parameters quantity $\beta > 0$ and hence the solution of the stationary problem is unique, however it is unstable since condition (4) is fulfilled. The computational simulation is based on the iterative process used above. The results of the computation of the value of the gap $g = q(q > 0)$ and of the contact

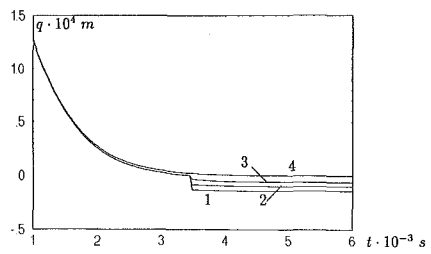


Fig. 6 Gap $g \equiv q(q > 0)$ and contact pressure $p \equiv p_0 q(q < 0)$ as functions of time for the set of values of g_0 and α_*: 1 – $1 \cdot 10^{-5}$, 2 – $1 \cdot 10^{-6}$, 3 – $5 \cdot 10^{-7}$ Pa^{-1}; 4 – $\alpha_* = 1 \cdot 10^{-5}$ Pa^{-1}, and $g_0 = 5.001 \cdot 10^{-4}$ m

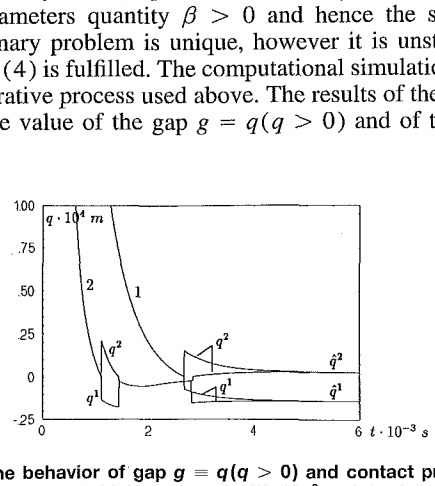


Fig. 8 Time behavior of gap $g = q(q > 0)$ and contact pressure $p = -p_0 q(q < 0)$ for the initial gap $g_0 = 2.127 \cdot 10^{-3}$ and the following values of the thermal perturbation $\vartheta_n(1 + \psi_n)$: $\vartheta_1 = 0°C$, $\vartheta_2 = 350°C$, $\vartheta_2\psi_2$: 1 – 0°C, 2 – 200°C

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

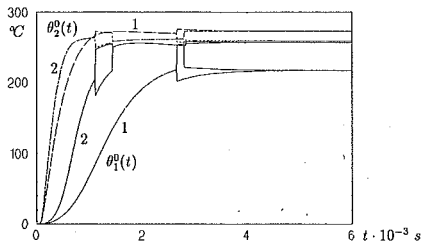


Fig. 9 Temperature of thermally contacting planes versus time for the initial gap $g_0 = 2.217 \cdot 10^{-3}$ and the following values of thermal perturbation $\vartheta_n(1 + \psi_n)$: $\vartheta_1 = 0°C$, $\vartheta_2 = 350°C$, $\vartheta_2\psi_2$: 1 – –0°C, 2 – –200°C

pressure $p = -p_0 q (q < 0)$ are presented in Fig. 8 by the curve marked "1". In Fig. 9 we have shown, by curves marked "1", the time behavior of the temperature of the contacting planes. Up to time $t_* = 2670$ s the iterative process converges to the unique, non-stationary solution of the considered problem. For the value $t_i = 2660$ s the sequence of the values of iterative process is as follows:

$$q_i^{(1)} = 4.015 \cdot 10^{-8}, \quad q_i^{(2)} = 2.804 \cdot 10^{-7},$$
$$q_i^{(3)} = 1.562 \cdot 10^{-7}, \quad q_i^{(4)} = 2.204 \cdot 10^{-8},$$
$$q_i^{(5)} = 1.872 \cdot 10^{-7}, \quad q_i^{(6)} = 2.044 \cdot 10^{-7},$$
$$q_i^{(7)} = 1.955 \cdot 10^{-7}, \quad q_i^{(8)} = 2.001 \cdot 10^{-7},$$
$$q_i^{(9)} = 1.977 \cdot 10^{-7}, \quad q_i^{(10)} = 1.989 \cdot 10^{-7},$$
$$q_i^{(11)} = 1.983 \cdot 10^{-7}, \quad q_i^{(12)} = 1.987 \cdot 10^{-7},$$
$$q_i^{(13)} = 1.985 \cdot 10^{-7}.$$

The iterative process ends with $j_i = 13$ for which the condition of the iterative process completion is fulfilled. The sequence $q_i^{(j)}$ converges to the value $q_i = q(2660)$ which is a fixed point of the mapping $Q(q, t_i)$ (see Eqs. (36) and (37)).

For $t > t_*$ the iterational sequence becomes divergent. The fixed point of mapping $Q(q, t_i)$ becomes unstable. For example the sequence of the values of the iterative process for $t_i = 2680$ s reads as follows:

$$q_i^{(1)} = -2.505 \cdot 10^{-7}, \quad q_i^{(2)} = 5.023 \cdot 10^{-6},$$
$$q_i^{(3)} = -2.679 \cdot 10^{-6}, \quad q_i^{(4)} = 1.332 \cdot 10^{-5},$$
$$q_i^{(5)} = -6.586 \cdot 10^{-6}, \quad q_i^{(6)} = 1.479 \cdot 10^{-5},$$
$$q_i^{(7)} = -7.243 \cdot 10^{-6}, \quad q_i^{(8)} = 1.489 \cdot 10^{-5},$$
$$q_i^{(9)} = -7.288 \cdot 10^{-6}, \quad q_i^{(10)} = 1.490 \cdot 10^{-5},$$
$$q_i^{(11)} = -7.291 \cdot 10^{-6}.$$

The sequence $q_i^{(2j-1)}$ converges to $q^1(2680)$ while the sequence $q_i^{(2j)}$ tends to value $q^2(2680)$ which constitute the stable fixed points of the mapping $Q(Q, t_i)$.

Thus when the system in the conditions of the decreasing gap becomes stationary and unstable, the solution becomes divergent.

Now let us consider the case when the system becomes likewise stationary and unstable, however, in the conditions of the decreasing contact pressure. We assume the following excitation of the system: $\vartheta_1 = 0°C$, $\vartheta_2 = 350°C$, $\vartheta_2\psi_2 = 200°C$, $h^2 = 2.5 \cdot 10^{-3} s^{-1}$. The results of the computation have been presented in Figs. 8 and 9 by curves marked "2". For $t \in (0, 1100]$ the nonstationary solution is unique, for $t \in [1110, 1430]$ the iterational process becomes divergent, for $t \in [1440, 2830]$ the iteration sequence is divergent while for $t \in [1440, 2830]$ the nonstationary solution becomes unique and tends to the unstable stationary one. For $t \geq 2840$ s the iteration sequence is again divergent. The branches of the solutions $q^1(t)$ and $q^2(t)$ tend to the values $\hat{q}^1$ and $\hat{q}^2$ which constitute the fixed points of the mapping $\hat{Q}(\hat{Q}(\hat{q}))$, see Fig. 8.

3 Conclusions

We have envisaged the transient thermoelastic process in a system consisting of two layers (or rods) in a state of an imperfect contact. The theoretical model is exactly that posed in Barber and Zhang (1988). It has been shown that in the regions where there exist the unique and stable stationary states, the transient process converges to a steady-state. In the regions of uniqueness and nonstability of the stationary state the transient state, for the considered thermal excitation, is not unique and has two branching solutions. On the other hand, in the case of nonuniqueness of the stationary state, the transient process converges, depending on the character of excitation, to the first or the second of stable stationary states. We also have shown that depending on the character of the thermal perturbation the unique and nonunique solutions alternate. It is interesting to note the interdependence between the conditions for the existence of the stable fixed point, and the investigations on the stability of various stationary states (compare Barber, Dundurs, and Comninou (1979) and Barber and Zhang (1988)).

We haven't included in this paper the investigations of the effect of different kinds of the thermal excitation of the system and the influence of the smoothness of the functions of the thermal resistance on the properties of the solutions. A number of problems remain open for further investigations. Within the considered mathematical model it would be necessary to find the general conditions, such that the solution was unique, and the conditions when there exist multiple solutions. In this paper we have envisaged characteristic examples. We have found that the solution could be unique, next doubled, again unique, and again branching. In the discussed model we have made use of the simplifying assumptions, namely the coupling term in the equation of the thermal conductivity, and the inertia term in generalized Navier's equation have been neglected. The justification was that the process was slow and both the terms, in classical problems, were negligible. We have assumed that this holds true also in the problem considered. However, at points when the solution changes its character from the unique to doubled solution, i.e., at the branching points we have found the rapid change of the field of temperature, consequently the assumption that the inertia term is negligible can be questioned. Then we would have to consider a more complicated system of partial differential equations of hyperbolic-parabolic type.

The statement about the existence of the stationary solutions results from the construction of the stationary solutions, Fig. 2. The statements about stability and nonstability of steady-state solutions have been discussed by Barber and Zhang (1988). The conditions of nonstability of intermediate steady-state and nonunique solution coincide with the conditions (43) (in Olesiak and Pyryev (1996)) of the nonstability of the fixed point of the mapping. The condition of the nonstability of the unique steady-state solution (44) (also in Olesiak and Pyryev, 1996) may serve as a condition for choosing the model parameters such that the iteration sequence is divergent while the subsequences converge. In the theory of nonlinear equations, for example in hydromechanics such conditions are called the conditions of the turbulence initiation. In the dynamical systems they correspond to the conditions of the initiation of chaotic motion.

Acknowledgment

The postdoctoral scholarship of the Ministry of Education of Poland for one of the authors (Yu. A. P.) is gratefully acknowledged. This support enabled him to work during the academic year 1992–1993 at the University of Warsaw.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

Barber, J. R., Dundurs, J., and Comninou, Maria, 1979, "Stability Considerations in Thermoelastic Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 871–874.

Barber, J. R., and Zhang, R., 1988, "Transient Behaviour and Stability for the Thermoelastic Contact of Two Rods of Dissimilar Materials," *International Journal of Mechanical Science,* Vol. 30, pp. 691–704.

Olesiak, Z. S., Pyryev, Yu. A., 1995, "Transient Response in a One-Dimensional Mode of Thermoelastic Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 575–581.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Higher-Order Theory for Vibration of Doubly Curved Shallow Shells

K. M. Liew
Division of Engineering Mechanics,
School of Mechanical and
Production Engineering,
Nanyang Technological University,
Singapore 639798
Mem. ASME

C. W. Lim
Department of Civil Engineering,
The University of Queensland,
Brisbane, Queensland 4072,
Australia

A higher-order shear deformation theory is presented for vibration analysis of thick, doubly curved shallow shells. An orthogonal curvilinear coordinate system is employed to arrive at the strain components. A third-order displacement field in transverse coordinate is adopted. Though no transverse normal stress is assumed, the theory accounts for cubic distribution of the transverse shear strains through the shell thickness in contrast with existing parabolic shear distribution. The unsymmetric shear distribution is a physical consequence of the presence of shell curvatures where the stress and strain of a point above the mid-surface are different from its counterpart below the mid-surface. Imposing the vanishing of transverse shear strains on top and bottom surfaces, the rotation field is reduced from a six-degree to a two-degree system. The discrepancy between the existing and the present theories is highlighted.

1 Introduction

Based on thin shell theory, a comprehensive research through the early 1970s has been summarized by Leissa (1973). In view of the shortcomings of the past research, the free vibration of thin shallow shells has been further discussed and investigated by Narita and Leissa (1986). A variety of shallow shell structures has been examined recently by Lim and Liew (1994) and Liew and Lim (1994a, b).

The neglect of transverse shear strains in thin shell theory results in erroneous response of structures. It underpredicts deflections and overpredicts frequency response and buckling loads. Therefore, a more adequate theory needs to be developed to overcome the shortcomings of thin shell theory. Since the mid 1980s Reddy and his associates have pioneered a series of studies on moderately thick to thick shells. Using a higher-order theory, Reddy (1984) proposed a shear deformation theory for doubly curved laminated shells. Exact solutions were obtained for simply supported, doubly curved, cross-ply laminated shells under sinusoidal, uniformly distributed, and point loads. Using a similar approach, Reddy and Liu (1985) provided the Navier-type exact solutions for bending and vibration of orthotropic laminated cylindrical and spherical shells under simply supported boundary conditions.

To complement the development for thick shell analysis, we present herein a higher-order shear deformation theory for vibration of thick shallow shells. This study accounts for the transverse shear strains and rotations. The displacement field is represented by a third-order expression in the transverse coordinate. The proposed rotation field has six degrees-of-freedom. Employing the vanishing of transverse shear strains on top and bottom surfaces, the rotation field is reduced to a two-degree system. It results in a total of five degrees-of-freedom: three for orthogonal displacements and two for rotations.

This paper is also to highlight the difference of Reddy's strain components and displacement expressions comparing to the present formulation. Reddy and his co-workers (Reddy, 1984; Reddy and Liu, 1985) have omitted the Lamé parameters for the transverse shear strains which resulted in the elimination of two of the rotation degrees and therefore obtained fewer terms in the displacement expressions. Perhaps the most distinct discrepancy in dropping the Lamé parameters is the resultant parabolic transverse shear strains in contrast with the cubic transverse shear strains (Lim and Liew, 1995; Liew and Lim, 1995). The present cubic shear strain distribution is not even (unsymmetric) with respect to the mid-surface while the parabolic shear strain distribution otherwise. The significant discrepancies between both theories are reported.

2 Theoretical Formulation

Consider a homogeneous, isotropic, thick doubly curved shallow shell of arbitrary boundary conditions and with x-dimension a, y-dimension b, thickness h as shown in Fig. 1. The orthogonal curvilinear coordinate system used is ($\bar{x}$, $\bar{y}$, $\bar{z}$) and the principal radii of curvature at mid-surface ($\bar{z} = 0$) are denoted by $R_{\bar{x}}$ and $R_{\bar{y}}$, respectively. The primary objective of the paper is to derive a higher-order theory for the vibration of such a shallow shell.

The formulation assumes the higher-order shear deformable displacement field as follows:

$$u = \left(1 + \frac{\bar{z}}{R_{\bar{x}}}\right) u_o + \bar{z}\theta_1^u + \bar{z}^2\theta_2^u + \bar{z}^3\theta_3^u \tag{1a}$$

$$v = \left(1 + \frac{\bar{z}}{R_{\bar{y}}}\right) v_o + \bar{z}\theta_1^v + \bar{z}^2\theta_2^v + \bar{z}^3\theta_3^v \tag{1b}$$

$$w = w_o \tag{1c}$$

where (u, v, w) are the displacements of an arbitrary point along the $\bar{x}$-, $\bar{y}$-, $\bar{z}$-directions and (u_o, v_o, w_o) are the corresponding displacements at mid-surface. The rotation parameters of a point are θ_i^u and θ_i^v $(i = 1, 2, 3)$.

Assuming no transverse normal strain, the strain components are

$$\epsilon_{\bar{x}} = \frac{1}{1 + \bar{z}/R_{\bar{x}}}\left(\frac{\partial u}{\partial \bar{x}} + \frac{w}{R_{\bar{x}}}\right) \tag{2a}$$

$$\epsilon_{\bar{y}} = \frac{1}{1 + \bar{z}/R_{\bar{y}}}\left(\frac{\partial v}{\partial \bar{y}} + \frac{w}{R_{\bar{y}}}\right) \tag{2b}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Sept. 28, 1994; final revision, June 18, 1995. Associate Technical Editor: J. N. Reddy.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

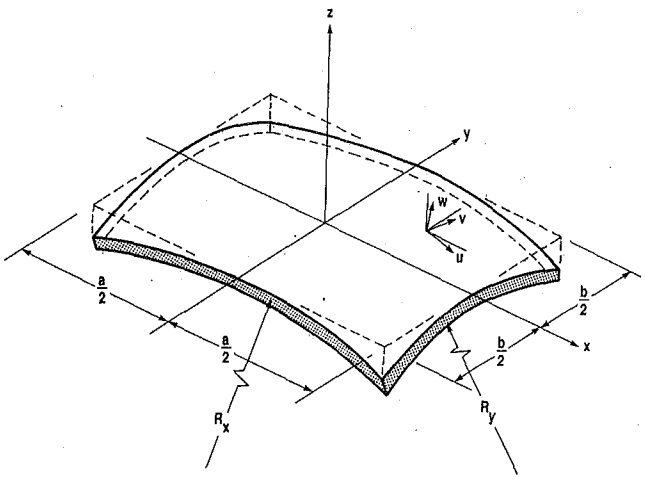


Fig. 1 Geometry of the thick cylindrical shell of rectangular planform

$$\epsilon_{\bar{z}} = 0 \tag{2c}$$

$$\gamma_{\bar{y}\bar{z}} = \frac{1}{1 + \bar{z}/R_{\bar{y}}}\left(\frac{\partial w}{\partial \bar{y}} - \frac{v}{R_{\bar{y}}}\right) + \frac{\partial v}{\partial \bar{z}} \tag{2d}$$

$$\gamma_{\bar{x}\bar{z}} = \frac{1}{1 + \bar{z}/R_{\bar{x}}}\left(\frac{\partial w}{\partial \bar{x}} - \frac{u}{R_{\bar{x}}}\right) + \frac{\partial u}{\partial \bar{z}} \tag{2e}$$

and

$$\gamma_{\bar{x}\bar{y}} = \frac{1}{1 + \bar{z}/R_{\bar{y}}}\frac{\partial u}{\partial \bar{y}} + \frac{1}{1 + \bar{z}/R_{\bar{x}}}\frac{\partial v}{\partial \bar{x}}. \tag{2f}$$

Substituting u, v, and w from Eqs. (1*a-c*) into Eqs. (2*a-f*) and using the vanishing of transverse shear strains on the shell top and bottom surfaces,

$$\gamma_{\bar{y}\bar{z}}(\bar{x}, \bar{y}, \pm h/2) = 0; \quad \gamma_{\bar{x}\bar{z}}(\bar{x}, \bar{y}, \pm h/2) = 0, \tag{3a, b}$$

result in the following expressions for the rotations:

$$\theta_2^u = \frac{1}{3R_{\bar{x}}}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{4a}$$

$$\theta_3^u = -\frac{4}{3h^2}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{4b}$$

$$\theta_2^v = \frac{1}{3R_{\bar{y}}}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) \tag{4c}$$

$$\theta_3^v = -\frac{4}{3h^2}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right). \tag{4d}$$

Subsequently the displacement field in Eqs. (1*a-c*) becomes

$$u = u_o + \bar{z}\left(\frac{u_o}{R_{\bar{x}}} + \theta_1^u\right) + \underline{\frac{\bar{z}^2}{3R_{\bar{x}}}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right)} - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{5a}$$

$$v = v_o + \bar{z}\left(\frac{v_o}{R_{\bar{y}}} + \theta_1^v\right) + \underline{\frac{\bar{z}^2}{3R_{\bar{y}}}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right)} - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) \tag{5b}$$

$$w = w_o. \tag{5c}$$

In the above derivation, terms of order $O(1/R_{\bar{x}}^2)$ and $O(1/R_{\bar{y}}^2)$ are dropped in view of the shallowness of the shell.

It is worth noting that Eqs. (2*d-e*) above are different from the expressions reported by Reddy and Liu (1985). They have omitted the Lamé parameters $(1 + \bar{z}/R_{\bar{x}})$ and $(1 + \bar{z}/R_{\bar{y}})$ underlined in Eqs. (2*d-e*) in their formulation for the transverse strains which results in $\theta_2^u = \theta_2^v = 0$ in Eqs. (1*a-b*) and consequently the absence of the underlined terms involving $\bar{z}^2$ in Eqs. (5*a-b*).

Using the above displacement field, the strain components in Eqs. (2*a-f*) can be expressed as

$$\epsilon_{\bar{x}} = \frac{\partial u_o}{\partial \bar{x}} + \frac{w_o}{R_{\bar{x}}} + \bar{z}\frac{\partial \theta_1^u}{\partial \bar{x}} + \frac{\bar{z}^2}{3R_{\bar{x}}}\left(\frac{\partial^2 w_o}{\partial \bar{x}^2} - 2\frac{\partial \theta_1^u}{\partial \bar{x}}\right) - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial^2 w_o}{\partial \bar{x}^2} + \frac{\partial \theta_1^u}{\partial \bar{x}}\right) \tag{6a}$$

$$\epsilon_{\bar{y}} = \frac{\partial v_o}{\partial \bar{y}} + \frac{w_o}{R_{\bar{y}}} + \bar{z}\frac{\partial \theta_1^v}{\partial \bar{y}} + \frac{\bar{z}^2}{3R_{\bar{y}}}\left(\frac{\partial^2 w_o}{\partial \bar{y}^2} - 2\frac{\partial \theta_1^v}{\partial \bar{y}}\right) - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial^2 w_o}{\partial \bar{y}^2} + \frac{\partial \theta_1^v}{\partial \bar{y}}\right) \tag{6b}$$

$$\epsilon_{\bar{z}} = 0 \tag{6c}$$

$$\gamma_{\bar{y}\bar{z}} = \frac{\partial w_o}{\partial \bar{y}} + \theta_1^v - \frac{\bar{z}}{3R_{\bar{y}}}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) - \frac{4\bar{z}^2}{h^2}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) + \frac{4\bar{z}^3}{3h^2 R_{\bar{y}}}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) \tag{6d}$$

$$\gamma_{\bar{x}\bar{z}} = \frac{\partial w_o}{\partial \bar{x}} + \theta_1^u - \frac{\bar{z}}{3R_{\bar{x}}}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) - \frac{4\bar{z}^2}{h^2}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) + \frac{4\bar{z}^3}{3h^2 R_{\bar{x}}}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{6e}$$

$$\gamma_{\bar{x}\bar{y}} = \frac{\partial u_o}{\partial \bar{y}} + \frac{\partial v_o}{\partial \bar{x}} + \bar{z}\left[\left(\frac{1}{R_{\bar{y}}} - \frac{1}{R_{\bar{x}}}\right)\left(\frac{\partial v_o}{\partial \bar{x}} - \frac{\partial u_o}{\partial \bar{y}}\right) + \frac{\partial \theta_1^u}{\partial \bar{y}} + \frac{\partial \theta_1^v}{\partial \bar{x}}\right] + \bar{z}^2\left[\left(\frac{1}{3R_{\bar{x}}} - \frac{1}{R_{\bar{y}}}\right)\frac{\partial \theta_1^u}{\partial \bar{y}} + \left(\frac{1}{3R_{\bar{y}}} - \frac{1}{R_{\bar{x}}}\right)\frac{\partial \theta_1^v}{\partial \bar{x}} + \frac{1}{3}\left(\frac{1}{R_{\bar{x}}} + \frac{1}{R_{\bar{y}}}\right)\frac{\partial^2 w_o}{\partial \bar{x}\partial \bar{y}}\right] - \frac{4\bar{z}^3}{3h^2}\left[2\frac{\partial^2 w_o}{\partial \bar{x}\partial \bar{y}} + \frac{\partial \theta_1^u}{\partial \bar{y}} + \frac{\partial \theta_1^v}{\partial \bar{x}}\right] + \frac{4\bar{z}^4}{3h^2}\left[\left(\frac{1}{R_{\bar{x}}} + \frac{1}{R_{\bar{y}}}\right)\frac{\partial^2 w_o}{\partial \bar{x}\partial \bar{y}} + \frac{1}{R_{\bar{y}}}\frac{\partial \theta_1^u}{\partial \bar{y}} + \frac{1}{R_{\bar{x}}}\frac{\partial \theta_1^v}{\partial \bar{x}}\right]. \tag{6f}$$

It is noted that the transverse strains, $\gamma_{\bar{y}\bar{z}}$ and $\gamma_{\bar{x}\bar{z}}$, are cubic expressions of the transverse coordinate $\bar{z}$ and they are not even functions (symmetric functions) with respect to mid-surface ($\bar{z} = 0$). It is, in fact, a natural consequence of having unsymmetric strain distribution through the shell thickness owing to the presence of curvatures where the stress and strain of a point above the mid-surface are different from its counterpart below the mid-surface.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

3 Derivation of Energy Functional

If no transverse normal stress (σ_z) is considered, the strain energy of the shell is

$$\mathcal{U} = \frac{E}{2(1-\nu^2)} \iiint_V \left[\epsilon_x^2 + 2\nu\epsilon_x\epsilon_y + \epsilon_y^2 + \frac{1-\nu}{2}(\gamma_{yz}^2 + \gamma_{xz}^2 + \gamma_{xy}^2) \right] dV \quad (7)$$

and the kinetic energy is

$$\mathcal{T} = \frac{\rho}{2} \iiint_V \left[\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2 + \left(\frac{\partial w}{\partial t}\right)^2 \right] dV \quad (8)$$

where V is the volume of shell and ρ is the mass density per unit volume.

The displacement field can be approximated by sinusoidal functions as follows:

$$(u_o, v_o, w_o, \theta_1^u, \theta_1^v)(\bar{x}, \bar{y}, t) = (U_o, V_o, W_o, \Theta_1^u, \Theta_1^v)(\bar{x}, \bar{y}) \sin \omega t. \quad (9)$$

Substituting Eqs. (6*a-f*) and (9) into Eqs. (7) and (8), the maximum strain and kinetic energies in a vibratory cycle are

$$\mathcal{U}_{\max} = \frac{D}{2} \iint_{\bar{A}} \left\{ \frac{12}{h^2} \left[\left(\frac{\partial U_o}{\partial \bar{x}}\right)^2 + \frac{2W_o}{R_{\bar{x}}}\frac{\partial U_o}{\partial \bar{x}} + \left(\frac{W_o}{R_{\bar{x}}}\right)^2 + \left(\frac{\partial V_o}{\partial \bar{y}}\right)^2 + \frac{2W_o}{R_{\bar{y}}}\frac{\partial V_o}{\partial \bar{y}} + \left(\frac{W_o}{R_{\bar{y}}}\right)^2 \right.\right.$$
$$+ 2\nu\left(\frac{\partial U_o}{\partial \bar{x}}\frac{\partial V_o}{\partial \bar{y}} + \frac{\partial U_o}{\partial \bar{x}}\frac{W_o}{R_{\bar{y}}} + \frac{\partial V_o}{\partial \bar{y}}\frac{W_o}{R_{\bar{x}}} + \frac{W_o^2}{R_{\bar{x}}R_{\bar{y}}}\right) + \frac{1-\nu}{2}\left(\frac{\partial U_o}{\partial \bar{y}}\right)^2 + (1-\nu)\frac{\partial U_o}{\partial \bar{y}}\frac{\partial V_o}{\partial \bar{x}} + \frac{1-\nu}{2}\left(\frac{\partial V_o}{\partial \bar{x}}\right)^2 \Big]$$
$$+ \frac{16(1-\nu)}{5h^2}\left[\left(\frac{\partial W_o}{\partial \bar{y}}\right)^2 + 2\frac{\partial W_o}{\partial \bar{y}}\Theta_1^v + (\Theta_1^v)^2 + \left(\frac{\partial W_o}{\partial \bar{x}}\right)^2 + 2\frac{\partial W_o}{\partial \bar{x}}\Theta_1^u + (\Theta_1^u)^2 \right] + \frac{2}{3R_{\bar{x}}}\frac{\partial U_o}{\partial \bar{x}}\frac{\partial^2 W_o}{\partial \bar{x}^2}$$
$$- \frac{4}{3R_{\bar{x}}}\frac{\partial U_o}{\partial \bar{x}}\frac{\partial \Theta_1^u}{\partial \bar{x}} + \frac{68}{105}\left(\frac{\partial \Theta_1^u}{\partial \bar{x}}\right)^2 - \frac{32}{105}\frac{\partial^2 W_o}{\partial \bar{x}^2}\frac{\partial \Theta_1^u}{\partial \bar{x}} + \frac{1}{21}\left(\frac{\partial^2 W_o}{\partial \bar{x}^2}\right)^2 + \frac{2}{3R_{\bar{y}}}\frac{\partial V_o}{\partial \bar{y}}\frac{\partial^2 W_o}{\partial \bar{y}^2} - \frac{4}{3R_{\bar{y}}}\frac{\partial V_o}{\partial \bar{y}}\frac{\partial \Theta_1^v}{\partial \bar{y}}$$
$$+ \frac{68}{105}\left(\frac{\partial \Theta_1^v}{\partial \bar{y}}\right)^2 - \frac{32}{105}\frac{\partial^2 W_o}{\partial \bar{y}^2}\frac{\partial \Theta_1^v}{\partial \bar{y}} + \frac{1}{21}\left(\frac{\partial^2 W_o}{\partial \bar{y}^2}\right)^2 + 2\nu\left[\frac{68}{105}\frac{\partial \Theta_1^u}{\partial \bar{x}}\frac{\partial \Theta_1^v}{\partial \bar{y}} + \frac{1}{3R_{\bar{y}}}\frac{\partial U_o}{\partial \bar{x}}\frac{\partial^2 W_o}{\partial \bar{y}^2} - \frac{2}{3R_{\bar{y}}}\frac{\partial U_o}{\partial \bar{x}}\frac{\partial \Theta_1^v}{\partial \bar{y}} \right.$$
$$+ \frac{1}{3R_{\bar{x}}}\frac{\partial V_o}{\partial \bar{y}}\frac{\partial^2 W_o}{\partial \bar{x}^2} - \frac{2}{3R_{\bar{x}}}\frac{\partial V_o}{\partial \bar{y}}\frac{\partial \Theta_1^u}{\partial \bar{x}} - \frac{16}{105}\frac{\partial^2 W_o}{\partial \bar{x}^2}\frac{\partial \Theta_1^v}{\partial \bar{y}} - \frac{16}{105}\frac{\partial^2 W_o}{\partial \bar{y}^2}\frac{\partial \Theta_1^v}{\partial \bar{x}} + \frac{1}{21}\frac{\partial^2 W_o}{\partial \bar{x}^2}\frac{\partial^2 W_o}{\partial \bar{y}^2} \Big]$$
$$+ \frac{1-\nu}{2}\left[\left(\frac{34}{15R_{\bar{x}}} - \frac{16}{5R_{\bar{y}}}\right)\frac{\partial U_o}{\partial \bar{y}}\frac{\partial \Theta_1^u}{\partial \bar{y}} + \left(\frac{34}{15R_{\bar{y}}} - \frac{16}{5R_{\bar{x}}}\right)\frac{\partial V_o}{\partial \bar{x}}\frac{\partial \Theta_1^v}{\partial \bar{x}} + \left(\frac{28}{15R_{\bar{x}}} + \frac{4}{15R_{\bar{y}}}\right)\frac{\partial V_o}{\partial \bar{x}}\frac{\partial^2 W_o}{\partial \bar{x}\partial \bar{y}} \right.$$
$$+ \left(\frac{28}{15R_{\bar{y}}} + \frac{4}{15R_x}\right)\frac{\partial U_o}{\partial \bar{y}}\frac{\partial^2 W_o}{\partial \bar{x}\partial \bar{y}} - \frac{14}{15R_{\bar{x}}}\frac{\partial V_o}{\partial \bar{x}}\frac{\partial \Theta_1^u}{\partial \bar{y}} - \frac{14}{15R_{\bar{y}}}\frac{\partial U_o}{\partial \bar{y}}\frac{\partial \Theta_1^v}{\partial \bar{x}} + \frac{68}{105}\left(\frac{\partial \Theta_1^u}{\partial \bar{y}}\right)^2 + \frac{136}{105}\frac{\partial \Theta_1^u}{\partial \bar{y}}\frac{\partial \Theta_1^v}{\partial \bar{x}}$$
$$\left.\left. + \frac{68}{105}\left(\frac{\partial \Theta_1^v}{\partial \bar{x}}\right)^2 - \frac{64}{105}\frac{\partial^2 W_o}{\partial \bar{x}\partial \bar{y}}\frac{\partial \Theta_1^u}{\partial \bar{y}} - \frac{64}{105}\frac{\partial^2 W_o}{\partial \bar{x}\partial \bar{y}}\frac{\partial \Theta_1^v}{\partial \bar{x}} + \frac{4}{21}\left(\frac{\partial^2 W_o}{\partial \bar{x}\partial \bar{y}}\right)^2 \right] \right\} d\bar{x}d\bar{y} \quad (10)$$

and

$$\mathcal{T}_{\max} = \frac{\rho\omega^2}{2} \iint_{\bar{A}} \left\{ h[U_o^2 + V_o^2 + W_o^2] \right.$$
$$+ \frac{h^3}{12}\left[\frac{2U_o}{3R_{\bar{x}}}\frac{\partial W_o}{\partial \bar{x}} + \frac{8}{3R_{\bar{x}}}U_o\Theta_1^u + \left(\frac{U_o}{R_{\bar{x}}}\right)^2 + (\Theta_1^u)^2 \right.$$
$$+ \frac{2V_o}{3R_{\bar{y}}}\frac{\partial W_o}{\partial \bar{y}} + \frac{8}{3R_{\bar{y}}}V_o\Theta_1^v + \left(\frac{V_o}{R_{\bar{y}}}\right)^2 + (\Theta_1^v)^2 \Big]$$
$$- \frac{h^3}{30}\left[\frac{U_o}{R_{\bar{x}}}\frac{\partial W_o}{\partial \bar{x}} + \frac{U_o\Theta_1^u}{R_{\bar{x}}} + \frac{\partial W_o}{\partial \bar{x}}\Theta_1^u + (\Theta_1^u)^2 \right.$$
$$+ \frac{V_o}{R_{\bar{y}}}\frac{\partial W_o}{\partial \bar{y}} + \frac{V_o\Theta_1^v}{R_{\bar{y}}} + \frac{\partial W_o}{\partial \bar{y}}\Theta_1^v + (\Theta_1^v)^2 \Big]$$
$$+ \frac{h^3}{252}\left[\left(\frac{\partial W_o}{\partial \bar{x}}\right)^2 + 2\frac{\partial W_o}{\partial \bar{x}}\Theta_1^u + (\Theta_1^u)^2 \right.$$
$$\left.\left. + \left(\frac{\partial W_o}{\partial \bar{y}}\right)^2 + 2\frac{\partial W_o}{\partial \bar{y}}\Theta_1^v + (\Theta_1^v)^2 \right] \right\} d\bar{x}d\bar{y}, \quad (11)$$

where $\bar{A}$ is the area of the shell mid-surface and $\mathcal{U}_{\max}$ and $\mathcal{T}_{\max}$ are, respectively, the maximum strain energy and maximum kinetic energy in a vibratory cycle.

An energy functional defined as the difference of the maximum strain and kinetic energy is introduced here,

$$\Pi = \mathcal{U}_{\max} - \mathcal{T}_{\max}, \quad (12)$$

which is useful in deriving the subsequent governing eigenvalue equation.

4 Governing Eigenvalue Equation

The integrations of Π given in Eqs. (10)–(11) involve the entire area of the mid-surface $\bar{A}$. However, in view of shallow shell analysis in the present study, the integration is performed on the projected planform area, A, of the curvilinear mid-surface onto a horizontal base. The new Cartesian coordinate system is designated by (x, y, z). As a result, $(\bar{x}, \bar{y}, \bar{z})$ in all the equations

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

above can be replaced by (x, y, z). Introducing the nondimensional coordinate system as follows,

$$\xi = \frac{x}{a}; \quad \eta = \frac{y}{b} \tag{13a, b}$$

where a and b are the length and width of the shell as shown in Fig. 1.

The displacement and rotation amplitude functions $U_o(\xi, \eta)$, $V_o(\xi, \eta)$, $W_o(\xi, \eta)$, $\Theta_1^u(\xi, \eta)$, and $\Theta_1^v(\xi, \eta)$ can be approximated by sets of two-dimensional series as follows:

$$U_o(\xi, \eta) = \sum_{i=1}^{m} c_i^u \phi_i^u(\xi, \eta) \tag{14a}$$

$$V_o(\xi, \eta) = \sum_{i=1}^{m} c_i^v \phi_i^v(\xi, \eta) \tag{14b}$$

$$W_o(\xi, \eta) = \sum_{i=1}^{m} c_i^w \phi_i^w(\xi, \eta) \tag{14c}$$

$$\Theta_1^u(\xi, \eta) = \sum_{i=1}^{m} \bar{c}_i^u \Phi_i^u(\xi, \eta) \tag{14d}$$

$$\Theta_1^v(\xi, \eta) = \sum_{i=1}^{m} \bar{c}_i^v \Phi_i^v(\xi, \eta) \tag{14e}$$

in which c_i^u, c_i^v, c_i^w, $\bar{c}_i^u$, $\bar{c}_i^v$ are the unknown coefficients and ϕ_i^u, ϕ_i^v, ϕ_i^w, Φ_i^u, Φ_i^v are the corresponding shape functions to be discussed in due course. The minimization of this energy functional with respect to the coefficients according to the Ritz procedure

$$\frac{\partial \Pi}{\partial c_i^\alpha} = 0; \quad \alpha = u, v, \text{ and } w \tag{15a}$$

and

$$\frac{\partial \Pi}{\partial \bar{c}_i^\alpha} = 0; \quad \alpha = u \text{ and } v \tag{15b}$$

lead to the governing eigenvalue equation

$$(\mathbf{K} - \lambda^2 \mathbf{M})\{\mathbf{C}\} = \{\mathbf{0}\}, \tag{16}$$

where the eigenvalue or nondimensional vibration frequency is

$$\lambda = \omega ab\sqrt{\frac{\rho h}{D}} \quad \text{or} \quad \lambda' = \omega a\sqrt{\frac{\rho}{E}}. \tag{17a, b}$$

The numerical results presented in the subsequent sections are in λ' instead of λ, because the former is independent of shell thickness.

5 Method of Solution

The displacement and rotation fields at mid-surface denoted by u_o, v_o, w_o, Θ_1^u, and Θ_1^v can be approximated by Eqs. (14*a-e*). Their corresponding *pb*-2 functions ϕ^u, ϕ^v, ϕ^w, Φ^u, and Φ^v can be expressed as

$$\varphi_i^\alpha(\xi, \eta) = f_i(\xi, \eta)\varphi_b^\alpha - \sum_{j=1}^{i-1} \Xi_{ij}^\alpha \varphi_j^\alpha \tag{18}$$

where

$$\Xi_{ij}^\alpha = \frac{{}_1\Delta_{ij}^\alpha}{{}_2\Delta_j^\alpha} \tag{19a}$$

$${}_1\Delta_{ij}^\alpha = \iint_A f_i(\xi, \eta)\varphi_1^\alpha \varphi_j^\alpha d\xi d\eta \tag{19b}$$

$${}_2\Delta_j^\alpha = \iint_A (\varphi_j^\alpha)^2 d\xi d\eta \tag{19c}$$

Table 1 The two-dimensional functions for various symmetry classes

Symmetry Class	U_o	V_o	W_o	Θ_1^u	Θ_1^v
SS	F_2	F_3	F_1	F_2	F_3
SA	F_1	F_4	F_2	F_1	F_4
AS	F_4	F_1	F_3	F_4	F_1
AA	F_3	F_2	F_4	F_3	F_2

$$\varphi = \phi, \Phi \quad \text{and} \quad \alpha = u, v, w.$$

The vibration modes can be divided into various symmetry classes with respect to the xz and yz-planes. For a shell having symmetric boundary conditions such as a hard-simply supported or fully clamped shell, the vibration modes can be classified into Symmetric-Symmetric (SS), Symmetric-Antisymmetric (SA), Antisymmetric-Symmetric (AS) and Antisymmetric-Antisymmetric (AA) classes. Therefore, the two-dimensional polynomial can be subdivided into odd and even functions in terms of powers of ξ and η:

$$\sum_{i=1}^{m} f_i(\xi, \eta) = F_1(\xi_e, \eta_e) + F_2(\xi_o, \eta_e) + F_3(\xi_e, \eta_o) + F_4(\xi_o, \eta_o) \tag{20}$$

where

$$F_1(\xi_e, \eta_e) = \sum_{q=0,2,4,\ldots}^{p} \sum_{i=0,2,4,\ldots}^{q} \xi^{q-i}\eta^i \tag{21a}$$

$$F_2(\xi_o, \eta_e) = \sum_{q=1,3,5,\ldots}^{p} \sum_{i=0,2,4,\ldots}^{q} \xi^{q-i}\eta^i \tag{21b}$$

$$F_3(\xi_e, \eta_o) = \sum_{q=1,3,5,\ldots}^{p} \sum_{i=1,3,5,\ldots}^{q} \xi^{q-i}\eta^i \tag{21c}$$

$$F_4(\xi_o, \eta_o) = \sum_{q=2,4,6,\ldots}^{p} \sum_{i=1,3,5,\ldots}^{q} \xi^{q-i}\eta^i. \tag{21d}$$

The degree of polynomial p is related to the number of terms m by

$$m = \frac{(p+1)(p+2)}{2}. \tag{22}$$

The two-dimensional functions of the displacement and rotation fields for various symmetry classes are presented in Table 1.

The basic function φ_b^α ($\varphi = \phi$, Φ and $\alpha = u, v, w$) are defined by the products of the equations of continuous piecewise boundary geometries of the shell planform each raised to an appropriate power that corresponds to the types of boundary constraints imposed on the shell, i.e.,

$$\phi_b^u(\xi, \eta) = \prod_{i=1}^{n_e} [\Upsilon_i(\xi, \eta)]^{\gamma_i^u}$$

$$\bullet\ \gamma_i^u = \begin{cases} 0 & \text{free; or soft simple support;} \\ & \text{or hard simple support } (y\text{-direction}) \\ 1 & \text{hard simple support } (x\text{-direction}); \\ & \text{or clamped.} \end{cases} \tag{23a}$$

$$\phi_b^v(\xi, \eta) = \prod_{i=1}^{n_e} [\Upsilon_i(\xi, \eta)]^{\gamma_i^v}$$

$$\bullet\ \gamma_i^v = \begin{cases} 0 & \text{free; or soft simple support;} \\ & \text{or hard simple support } (x\text{-direction}) \\ 1 & \text{hard simple support } (y\text{-direction}); \\ & \text{or clamped.} \end{cases} \tag{23b}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 2 Comparison of λ' for a thick spherical shallow shell with ν = 0.3, a/b = 1.0 and R_y/R_x = 1.0

$\frac{b}{R_y}$	$\frac{h}{b}$	Ref.	Mode Sequence Number								
			SS-1	SS-2	SS-3	SA-1 (AS-1)	SA-2 (AS-2)	SA-3 (AS-3)	AA-1	AA-2	AA-3
(a) SSSS											
0.5	0.1	‡	0.75315	2.6148	2.6148	1.4559	1.9481	3.2615	2.1631	3.8962	3.8962
		§	0.76049	2.6239	2.6239	1.4649	1.9481	3.2706	2.1723	3.8962	3.8962
	0.5	‡	1.9041	2.7447	4.6526	1.9420	3.4695	4.3318	3.8769	3.8769	4.6347
		§	1.9453	2.7308	4.5808	1.9384	3.5060	4.2851	3.8487	3.8487	4.6711
0.7	0.1	‡	0.88484	2.6530	2.6530	1.5271	1.9479	3.2912	2.2100	3.8958	3.8958
		§	0.89693	2.6705	2.6705	1.5438	1.9479	3.3090	2.2276	3.8957	3.8957
	0.5	‡	1.9442	2.7347	4.6500	1.9361	3.4910	4.3090	3.8585	3.8585	4.6507
		§	2.0214	2.7083	4.5206	1.9290	3.5610	4.2242	3.8064	3.8064	4.7214
(b) CCCC											
0.5	0.1	‡	1.2063	3.1517	3.1956	1.9446	3.7172	3.8317	2.6925	4.4382	5.1580
		§	1.2115	3.1592	3.2028	1.9515	3.7190	3.8375	2.7000	4.4381	5.1659
	0.5	‡	2.4717	5.6115	5.7153	3.6005	3.9270	5.5368	4.4137	4.9816	5.4175
		§	2.4916	5.6523	5.7427	3.6173	3.9000	5.5959	4.3672	5.0286	5.3185
0.7	0.1	‡	1.3815	3.1811	3.2467	1.9978	3.7195	3.8727	2.7362	4.4369	5.1606
		§	1.3905	3.1956	3.2544	2.0107	3.7234	3.8840	2.7506	4.4368	5.1743
	0.5	‡	2.5648	5.6050	5.7314	3.5422	3.9979	5.5694	4.3900	4.9760	5.4067
		§	2.6026	5.6757	5.7839	3.5546	3.9689	5.6815	4.3056	5.0616	5.2360

‡ HSDT: Independent calculations using expressions of Reddy and Liu (1985)
§ HSDT*: Present refined higher-order theory

$$\phi_b^w(\xi, \eta) = \prod_{i=1}^{n_e} [\Upsilon_i(\xi, \eta)]^{\gamma_i^w}$$

$$\bullet\ \gamma_i^w = \begin{cases} 0 & \text{free;} \\ 1 & \text{soft; hard simple supports;} \\ 2 & \text{clamped.} \end{cases} \tag{23c}$$

$$\Phi_b^u(\xi, \eta) = \prod_{i=1}^{n_e} [\Upsilon_i(\xi, \eta)]^{\Gamma_i^u}$$

$$\bullet\ \Gamma_i^u = \begin{cases} 0 & \text{free; soft simple support;} \\ & \quad \text{or hard simple support } (y\text{-direction}); \\ 1 & \text{hard simple support } (x\text{-direction}); \\ & \quad \text{or clamped.} \end{cases} \tag{23d}$$

$$\Phi_b^v(\xi, \eta) = \prod_{i=1}^{n_e} [\Upsilon_i(\xi, \eta)]^{\Gamma_i^v}$$

$$\bullet\ \Gamma_i^v = \begin{cases} 0 & \text{free; soft simple support;} \\ & \quad \text{or hard simple support } (x\text{-direction}); \\ 1 & \text{hard simple support } (y\text{-direction}); \\ & \quad \text{or clamped.} \end{cases} \tag{23e}$$

where n_e is the number of supporting edges and Υ_i is the boundary expression of the ith supporting edge.

6 Results and Discussion

Two types of boundary conditions are examined, namely, shells with simple hard supports (SSSS) and fully clamped (CCCC) boundaries. A convergence study on the eigenvalue of Eq. (16) is performed by increasing steadily the number of terms of the pb-2 shape function m for each symmetry class given by Eq. (20). Downward convergence of eigenvalue is observed indicating the existence of an upperbound eigenvalue. From this convergence study, it is concluded that a total of 40 terms for each degree-of-freedom or equivalently a determinant size of 40 × 5 in Eq. (16) is able to achieve convergent results. Thus, m = 40 is used throughout the present study.

A study is further carried out to examine the discrepancies between Reddy's higher-order shear deformation theory (HSDT) and the present higher-order shear deformation theory (HSDT*). The HSDT solutions are computed in a similar way following exactly identical displacement and strain fields presented by Reddy and Liu (1985):

$$u = u_o + \bar{z}\left(\frac{u_o}{R_{\bar{x}}} + \theta_1^u\right) - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{24a}$$

$$v = v_o + \bar{z}\left(\frac{v_o}{R_{\bar{y}}} + \theta_1^v\right) - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) \tag{24b}$$

$$w = w_o \tag{24c}$$

and

$$\epsilon_{\bar{x}} = \frac{\partial u_o}{\partial \bar{x}} + \frac{w_o}{R_{\bar{x}}} + \bar{z}\frac{\partial \theta_1^u}{\partial \bar{x}} - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial^2 w_o}{\partial \bar{x}^2} + \frac{\partial \theta_1^u}{\partial \bar{x}}\right) \tag{25a}$$

$$\epsilon_{\bar{y}} = \frac{\partial v_o}{\partial \bar{y}} + \frac{w_o}{R_{\bar{y}}} + \bar{z}\frac{\partial \theta_1^v}{\partial \bar{y}} - \frac{4\bar{z}^3}{3h^2}\left(\frac{\partial^2 w_o}{\partial \bar{y}^2} + \frac{\partial \theta_1^v}{\partial \bar{y}}\right) \tag{25b}$$

$$\epsilon_{\bar{z}} = 0 \tag{25c}$$

$$\gamma_{y\bar{z}} = \frac{\partial w_o}{\partial \bar{y}} + \theta_1^v - \frac{4\bar{z}^2}{h^2}\left(\frac{\partial w_o}{\partial \bar{y}} + \theta_1^v\right) \tag{25d}$$

$$\gamma_{x\bar{z}} = \frac{\partial w_o}{\partial \bar{x}} + \theta_1^u - \frac{4\bar{z}^2}{h^2}\left(\frac{\partial w_o}{\partial \bar{x}} + \theta_1^u\right) \tag{25e}$$

$$\gamma_{xy} = \frac{\partial u_o}{\partial \bar{y}} + \frac{\partial v_o}{\partial \bar{x}} + \bar{z}\left(\frac{\partial \theta_1^u}{\partial \bar{y}} + \frac{\partial \theta_1^v}{\partial \bar{x}}\right) - \frac{4\bar{z}^3}{3h^2}\left[2\frac{\partial^2 w_o}{\partial \bar{x}\partial \bar{y}} + \frac{\partial \theta_1^u}{\partial \bar{y}} + \frac{\partial \theta_1^v}{\partial \bar{x}}\right]. \tag{25f}$$

It is noticed that the in-plane displacement field in Eqs. (24*a-b*) is identical to Eqs. (5*a-b*) if the underlined terms are dropped. This is due to the neglect of Lamé parameters for the

Table 3 Frequency parameter λ' for a thick shallow shell with ν = 0.3 and a/b = 1.0

$\frac{R_y}{R_x}$	$\frac{b}{R_y}$	$\frac{h}{b}$	Mode Sequence Number								
			SS-1	SS-2	SS-3	SA-1 (AS-1)	SA-2 (AS-2)	SA-3 (AS-3)	AA-1	AA-2	AA-3
(a) SSSS											
-1.0	0.1	0.1	0.57614	2.5752	2.5752	1.3773	1.9483	3.2295	2.1117	3.8966	3.8966
		0.3	1.4116	2.7588	4.6419	1.9482	2.8717	4.3582	3.8963	3.8963	3.9942
		0.5	1.8555	2.7599	4.6396	1.9479	3.4443	4.3586	3.8947	3.8947	4.6125
	0.5	0.1	0.55750	2.5987	2.5987	1.3963	1.9481	3.2194	2.0931	3.8962	3.8962
		0.3	1.3592	2.8379	4.6403	1.9459	2.8531	4.3968	3.8878	3.8878	3.9380
		0.5	1.7711	2.8595	4.4724	1.9382	3.3934	4.3664	3.8460	3.8460	4.5050
1.0	0.1	0.1	0.58556	2.5762	2.5762	1.3802	1.9483	3.2316	2.1150	3.8966	3.8966
		0.3	1.4178	2.7552	4.6435	1.9482	2.8748	4.3561	3.8963	3.8963	3.9985
		0.5	1.8630	2.7544	4.6540	1.9479	3.4489	4.3536	3.8947	3.8947	4.6197
	0.5	0.1	0.76049	2.6239	2.6239	1.4649	1.9481	3.2706	2.1723	3.8962	3.8962
		0.3	1.5062	2.7505	4.6729	1.9459	2.9273	4.3441	3.8879	3.8879	4.0420
		0.5	1.9453	2.7308	4.5808	1.9384	3.5060	4.2851	3.8487	3.8487	4.6711
(b) CCCC											
-1.0	0.1	0.1	0.99239	3.1216	3.1499	1.8867	3.7269	3.7775	2.6457	4.4403	5.1383
		0.3	1.9939	4.9844	5.0528	3.3264	3.7286	5.8677	4.3772	4.4689	5.4358
		0.5	2.3724	5.6153	5.6978	3.7187	3.7956	5.5040	4.4346	4.9876	5.4353
	0.5	0.1	1.1265	3.1540	3.1785	1.9294	3.7295	3.7899	2.6519	4.4590	5.1307
		0.3	2.0685	5.0060	5.0621	3.3142	3.7690	5.7863	4.2146	4.6431	5.4273
		0.5	2.4388	5.6367	5.6881	3.6343	3.8825	5.5512	4.3338	5.0812	5.4161
1.0	0.1	0.1	0.99645	3.1218	3.1509	1.8876	3.7258	3.7801	2.6477	4.4394	5.1397
		0.3	1.9962	4.9846	5.0540	3.3264	3.7295	5.8734	4.4075	4.4390	5.4358
		0.5	2.3746	5.6160	5.6998	3.7175	3.7968	5.5059	4.4364	4.9849	5.4308
	0.5	0.1	1.2115	3.1592	3.2028	1.9515	3.7190	3.8375	2.7000	4.4381	5.1659
		0.3	2.1241	5.0098	5.0928	3.3146	3.7919	5.9065	4.4268	4.4409	5.4283
		0.5	2.4916	5.6523	5.7427	3.6173	3.9000	5.5959	4.3672	5.0286	5.3185

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

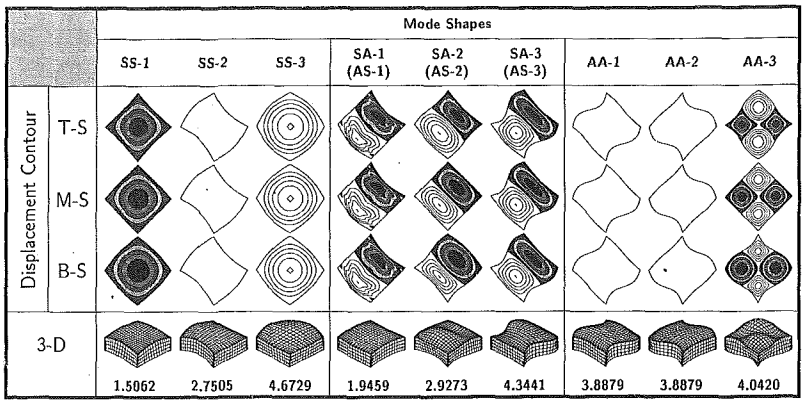


Fig. 2 Contour and three-dimensional vibration mode shapes of a doubly curved SSSS thick shallow shell with $\nu = 0.3$, $a/b = 1.0$, $b/R_y = 0.5$, $R_y/R_x = 1.0$ and $h/b = 0.3$ (T-S: top surface, M-S: mid-surface and B-S: bottom surface). Blank mode shapes are pure in-plane vibration modes.

transverse shear strains underlined in Eqs. (2*d-e*). The omission of these parameters results in the elimination of two of the rotation parameters, θ_2^u and θ_2^v, as in the formulation of the higher-order plate theory. This formulation yields parabolic transverse shear strains (Reddy and Liu 1985) which is symmetric with respect to the shell mid-surface. However, the present transverse shear strains given in Eqs. (6*d-e*) are cubic and unsymmetric expressions. The presence of unsymmetric transverse shear distribution is in fact a definite consequence of the presence of shell curvature.

Table 2 shows a comparison of solutions computed by the aforementioned formulations for a spherical shallow shell with square platform subject to SSSS and CCCC boundary conditions. The frequency parameters are arranged in various symmetry classes with reference to the *xz*- and *yz*-planes. Owing to the symmetry of shell geometries and boundary conditions, the SA and AS modes are identical. The solutions of HSDT and HSDT* diverges when h/b increases from 0.1 to 0.5.

Table 3 presents the nondimensional frequency parameters λ' of the SSSS and CCCC thick shells. The hyperbolic paraboloidal shells have negative Gaussian curvature ($R_y/R_x < 0$) whereas the spherical shells otherwise ($R_y/R_x > 0$). The hard simple support conditions expressed in Eqs. (23*a-e*) refers to boundaries where no rotation in the edge surfaces is allowed. However, rotations of boundaries are permitted in the planes normal to the edge surfaces. The fully clamped constraint implies that the edge surfaces are refrained from deflection and rotation. In this table, it is observed that an increase in shell thickness increases the vibration frequency and thus induces a shell with higher stiffness.

To visualize the vibrating behavior, illustrations of transverse displacement contour (w) and three-dimensional mode shapes is illustrated in Figs. 2 and 3 for the SSSS and CCCC shells. Note that the displacement amplitude in the three-dimensional modes is highly exaggerated but the deflection components u, v, and w are magnified proportionally. The contour plots for the top (T-S), mid (M-S) and bottom surfaces (B-S) are presented where the shaded regions represents regions with negative transverse displacement amplitude while the unshaded regions otherwise. The lines of demarcation are the nodal lines.

7 Closing Remarks

This paper presents a higher-order shear deformation theory for free vibrations of doubly curved shallow shells of arbitrary boundary conditions. The theory assumes no transverse normal stress. By employing the transverse shear-free conditions on the top and bottom surfaces of the shell, the six-degree rotation field can be reduced to a two-degree field. Furthermore, the analysis account for a cubic distribution (noneven terms) of the transverse shear strains through the shell thickness in contrast to the parabolic shear distribution (even terms) of Reddy and Liu (1985) in which the Lamé parameters for the transverse shear strain have been neglected. The unsymmetric strain distribution is a natural consequence due to the presence of shell curvature where the stress and strain of an arbitrary point above

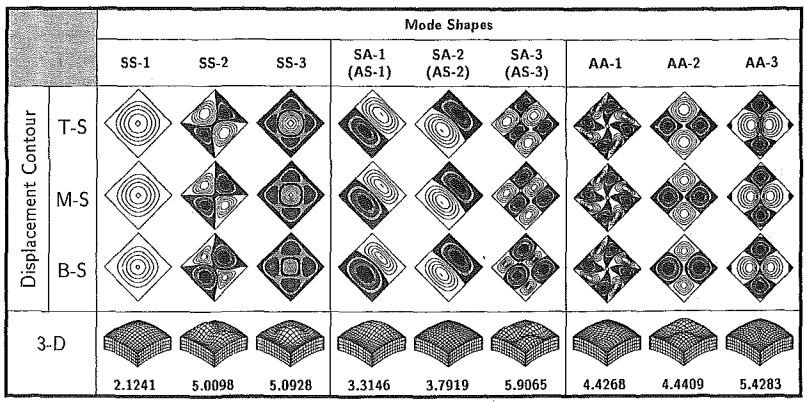


Fig. 3 Contour and three-dimensional vibration mode shapes of a doubly curved CCCC thick shallow shell with $\nu = 0.3$, $a/b = 1.0$, $b/R_y = 0.5$, $R_y/R_x = 1.0$ and $h/b = 0.3$ (T-S: top surface, M-S: mid-surface and B-S: bottom surface)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the shell mid-surface are different from its counterpart below the mid-surface.

References

Leissa, A. W., 1973, "Vibration of shells," NASA SP-288.

Liew, K. M., and Lim, C. W., 1994a, "Vibratory characteristics of cantilevered shallow shells of variable thickness," *American Institute of Aeronautics and Astronautics Journal,* Vol. 32, No. 2, pp. 387–396.

Liew, K. M., and Lim, C. W., 1994b, "Vibration of Perforated Doubly-Curved Shallow Shells with Rounded Corners," *International Journal of Solids and Structures,* Vol. 31, No. 11, pp. 1519–1536.

Liew, K. M., and Lim, C. W., 1995, "A higher-order theory for vibration analysis of curvilinear thick shallow shells with constrained boundaries," *Journal of Vibration and Control,* Vol. 1, pp. 15–39.

Lim, C. W., and Liew, K. M., 1994, "A *pb-2* Ritz formulation for flexural vibration of shallow cylindrical shells of rectangular planform," *Journal of Sound and Vibration,* Vol. 173, pp. 343–375.

Lim, C. W., and Liew, K. M., 1995, "A higher order theory for vibration of shear deformable cylindrical shallow shells," *International Journal of Mechanical Sciences,* Vol. 37, No. 3, pp. 277–295.

Narita, Y., and Leissa, A. W., 1986, "Vibrations of Completely Free Shallow Shells of Curvilinear Planform," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 647–651.

Reddy, J. N., 1984, "Exact Solutions of Moderately Thick Laminated Shells," *ASCE Journal of Engineering Mechanics,* Vol. 110, pp. 794–809.

Reddy, J. N., and Liu, C. F., 1985, "A Higher-order Shear Deformation Theory of Laminated Elastic Shells," *International Journal of Engineering Science,* Vol. 23, pp. 319–330.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Freezing Range Effect on Shell Growth Instability During Alloy Solidification

L. G. Hector, Jr.
Surface Technology Division,
Alcoa Technical Center-C,
100 Technical Drive,
Alcoa Center, PA 15069
Mem. ASME

Woo-Seung Kim
Mechanical Engineering Department,
Hanyang University,
Seoul 133-791, Korea

O. Richmond
Corporate Fellow,
Alcoa Technical Center,
100 Technical Drive,
Alcoa Center, PA 15069
Mem. ASME

During metal casting, nonuniform thermal strains due to small spatial perturbations in the cooling profile may lead to a condition of growth instability. This condition is one where the metal shell thickness becomes highly nonuniform due to the formation of microscopic air gaps along the mold surface. Documented experimental work with iron-carbon and aluminum alloy systems provides qualitative evidence that the freezing range width has an important influence on the macromorphology of the freezing front and hence the shell thickness. In this paper, we develop a thermomechanical model which points to a possible mechanism for growth instability during the solidification of a three-phase system where a small, spatially periodic cooling profile is superposed onto uniform cooling. This extends earlier work on two-phase (pure metal) systems. Temperature fields and shell growth are first calculated using perturbation theory, and then the associated thermal stresses and strains are determined from a hypoelastic constitutive law. The evolution of the shell/mold contact pressure beneath a shell thickness minimum is examined as a function of the freezing range width, and the onset of growth instability is assumed to occur as the contact pressure drops to zero indicating possible air gap nucleation. Under fixed cooling conditions, instability increases with freezing range width at small liquid metal pressures, decreases with freezing range width at somewhat larger liquid metal pressures, and disappears at still larger pressures.

Introduction

Small spatial perturbations in the heat extraction profile applied to the mold surface may lead to a nonuniform shell thickness and irregular contraction of the shell during casting. The contractions produce variations in the contact pressure that lead to intermittent air gap nucleation at the shell/mold interface. If the contact pressure decreases beneath the thinnest regions of the shell, heat extraction is diminished and further growth of the shell is locally inhibited. If the contact pressure simultaneously increases beneath the thickest regions of the shell, growth of these regions is further enhanced and the shell thickness becomes highly nonuniform. This condition can promote cellular undulation where the macromorphology of the freezing front and hence the shell thickness become highly irregular due to localized contractions beneath the thinnest regions of the shell. The resulting shell growth instability leads to surface and subsurface cracking, surface exudation, and a highly irregular microstructure.

Singh and Blazek (1974) and Cisse et al. (1971) conducted casting experiments that examined the influence of freezing range width on growth instability during alloy solidification. However, the issue as to whether an alloy with a larger freezing range might be more susceptible to growth instability as opposed to one with a narrower freezing range is not clearly resolved when comparing the results presented in these articles. For example, Singh and Blazek observed growth instability in continuously cast iron-carbon alloys and concluded that irregular growth of the metal shell is related to alloy composition. Nonuniform variations in shell thickness were most prominent for iron-carbon alloys with 0.1wt.%C. As the carbon content increased, nonuniform growth rapidly diminished. According to the phase diagram (see DeGarmo, 1988), that alloy with the largest freezing range has 0.1wt.%C. Singh and Blazek's work therefore seems to suggest that continuously cast alloys having the widest freezing ranges tend to be most susceptible to growth instability. On the other hand, Cisse et al. studied growth instability during static casting of aluminum alloys in cylindrical molds. They found that the extent of the freezing range and the temperature gradient at the shell/mold interface are the main variables influencing freezing front growth instability. Those alloys with the smallest freezing ranges exhibited the greatest growth instability, a result that seems to contradict that due to Singh and Blazek. It is therefore difficult to draw any substantial conclusions about the freezing range effect on shell growth instability from these articles.

Theoretical models of growth instability during directional solidification of pure metals have been advanced by Richmond et al. (1990), Li and Barber (1991), and Yigit and Barber (1992). In each of these models, a spatially periodic cooling profile leads to nonuniform thermal strains in the metal shell. Richmond et al. determined that a hypoelastic material with a negligible Stefan number has the potential for unstable growth since air gaps nucleate beneath the thinnest regions of the shell. These authors also discussed competing instability mechanisms in solidification, such as the constitutional supercooling and Mullins-Sekerka theories. Li and Barber developed a stress function approach wherein the thermal and mechanical fields are coupled along the mold interface through a pressure-dependent thermal contact resistance. Yigit and Barber found that materials with higher specific heats (or smaller diffusivities) may be less prone to growth instability. Their paper contains a photograph of a 3004 aluminum shell which shows a clear example of morphological instability during alloy solidification.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the International Mechanical Engineering Congress and Exposition, the Winter Annual Meeting of The ASME, Atlanta, GA, Nov. 17–22, 1996.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Jan. 4, 1995; final revision, Mar. 4, 1996. Associate Technical Editor: J. T. Jenkins.

Paper No. 96-WA/APM-2.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

In order to isolate the influence of freezing range width on growth instability of the metal shell, we shall consider the idealized case of a zero Stefan number alloy solidifying on a planar mold surface. The shell is assumed to deform according to a hypoelastic constitutive equation which is a rate formulation of the classical elastic constitutive law. The solidification process occurs in two stages. In Stage I, a uniform heat flux is applied to the mold surface on which rests a quiescent molten fluid at its liquidus temperature. A two-phase or mush region develops in which latent heat is continuously released and subsequently extracted through the mold surface. Stage I ends when a contiguous layer of solid nucleates along the mold surface. At the beginning of Stage II, a small sinusoidal perturbation, which is an idealized representation of stochastic variations in the cooling profile due to mold surface topography, interfacial oxides, and lubricants, is superposed onto the uniform cooling profile of Stage I and the solid shell begins to deform. A thermohypoelastic stress function approach is used to examine the evolution of the contact pressure beneath the thinnest regions of the shell as a function of the nominal freezing range width in Stage II. The analysis is terminated once the contact pressure drops to zero as the model cannot address the complexities associated with convection and radiation heat transfer in an air gap. Hence, the developments herein, like our earlier work on this subject, point to a plausible thermomechanical mechanism for the onset of growth instability, but do not address post instability conditions.

The Heat Conduction Problem

Stage I: Mush and Liquid. At the start of Stage I, a quiescent bath of molten fluid uniformly at the liquidus temperature, T_l, rests on a semi-infinite mold. As shown in Fig. 1, a mush region forms after a uniform cooling profile, Q_o, is applied to the mold surface. The width of the mush region relative to the mold surface is $a_o(t)$. The thermal field in the mush region, $T_m(z, t)$, is governed by

$$\frac{\partial^2 T_m(z,t)}{\partial z^2} + \frac{\rho L}{K_m}\frac{\partial f(z,t)}{\partial t} = 0; \quad 0 \le t \le t_o \tag{1}$$

subject to the following conditions:

$$K_m \frac{\partial T_m(0,t)}{\partial z} = Q_o; \quad T_m(0,t_o) = T_s^*; \quad a_o(0) = 0 \tag{2}$$

$$T_m(a_o,t) = T_l^*; \quad K_m \frac{\partial T_m(a_o,t)}{\partial z} = 0. \tag{3}$$

Equation (2*b*) states that the mold surface temperature drops to the solidus temperature at an unknown time t_o (to be determined in the analysis). At this time, a contiguous layer of solid has nucleated along the mold surface thereby signaling the end of Stage I and the beginning of Stage II, where solid shell deformation begins. Note that the solutal diffusion problem is neglected in both stages of the analysis and hence the model is based upon an idealized alloy system.

The solid fraction in the mush region, $f(z, t)$, is assumed to be

$$f(z,t) = \left(\frac{T_l^* - T_m(0,t)}{T_l^* - T_s^*}\right)\left(1 - \frac{z}{a_o}\right). \tag{4}$$

Linear approximations to the solid fraction distribution are clearly idealizations: more realistic forms of the solid fraction

Nomenclature

$a(x, t)$ = thickness of mush region
$a_o(t)$ = mean thickness of mush region
$a_1(t)$ = mush thickness perturbation
$b(x, t)$ = solid shell thickness
$b_o(t)$ = mean thickness of solid shell
$b_1(t)$ = solid shell thickness perturbation
$f(x, z, t)$ = solid fraction distribution in mush region
$f_o(z, t)$ = solid fraction corresponding to zeroth-order solution
$f_1(z, t)$ = solid fraction corresponding to first-order solution
f_{sb} = solid fraction at shell/mush boundary
$g(z)$ = function representing residual stress
$\bar{g}(z|\xi; t)$ = Green's function
m = wave number ($2\pi/\omega$)
p = hydrostatic pressure at shell/mush interface
t = time
t_o = time at which Stage II begins
$\tilde{u}$ = displacement field
w = tanh (mb_o)
x = spatial coordinate along mold surface
z = spatial coordinate normal to mold surface
$A = 2E\alpha/K_s$
$C_i(t)$ = coefficients of Airy stress function
$D_i(t)$ = functions defined by Eqs. (83)
E = elastic modulus
$F(\xi, t)$ = distributed source
K = thermal conductivity
$G(\eta)$ = dimensionless form of $g(z)$
L = latent heat
$P(x, t)$ = contact pressure at shell/mold interface during Stage II
$\bar{P}$ = dimensionless contact pressure
$P_o(t)$ = zeroth-order contact pressure
$\bar{P}_o(\eta)$ = dimensionless zeroth-order contact pressure
$P_l(t)$ = first-order contact pressure
$\bar{P}_l(\eta)$ = dimensionless first-order contact pressure
Q_o = uniform heat flux
Q_l = heat flux perturbation
$T(x, z, t)$ = temperature field
$T_m(0, t)$ = mush temperature at mold surface during Stage I
T_s^* = solidus temperature
T_l^* = liquidus temperature
α = thermal expansion coefficient
δ = delta function
η = dimensionless time variable
λ = mean width of freezing range during Stage II
$\bar{\lambda}$ = dimensionless mean width of freezing range during Stage II
ν = Poisson's ratio
ξ = source point in Green's function
ρ = density
σ_{ij} = total stress tensor
$\varphi(x, z, t)$ = displacement potential
ψ_i = functions associated with distributed source term in temperature solution
ω = wavelength of spatial perturbation in cooling profile
$\Phi(x, z, t)$ = Airy stress function

Subscripts

l = liquid region
m = mush region
s = solid shell
o = zeroth-order problem
1 = first-order problem

Superscripts

h = homogeneous solution
p = particular solution
r = residual stress
o = zeroth-order problem
1 = first-order problem

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

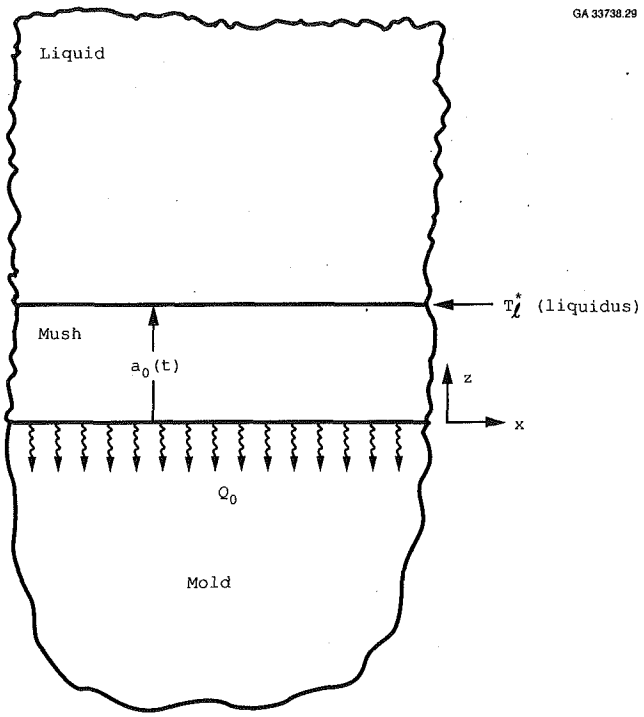


Fig. 1 Solidification system during Stage I

distribution in aluminum alloy systems (for example) are found in Bäckerud et al. (1986). The time derivative of Eq. (4) (for use in Eq. (1)) is written as

$$\dot{f}(z, t) = \psi_1(t) + z\psi_2(t) \tag{5}$$

where

$$\psi_1(t) = \dot{f}_o; \quad \psi_2(t) = \frac{f_o\dot{a}_o - \dot{f}_o a_o}{a_o^2}. \tag{6}$$

Note that the unknown function $f_o(t)$ is the solid fraction distribution along the mold surface. It is related to the mush temperature at the mold surface, $T_m(0, t)$, through

$$f_o(t) = \frac{T_l^* - T_m(0, t)}{T_l^* - T_s^*}. \tag{7}$$

The temperature field in the mush region is the sum of a homogeneous solution, T_m^h, which satisfies the boundary conditions but not Eq. (1), and a particular solution, T_m^p, which satisfies Eq. (1) but not all of Eqs. (2) and (3). Using the homogeneous form of Eq. (1) subject to Eqs. (2a) and (3a), we find

$$T_m^h(z, t) = T_l^* - \frac{Q_o a_o}{K_m}\left(1 - \frac{z}{a_o}\right) \tag{8}$$

where $a_o(t)$ is the unknown mush thickness.

We adopt the Green's function technique (see Stakgold, 1967) for the particular solution to remain consistent with the solution procedure to be used for Stage II. The temperature corresponding to the particular solution, which is governed by Eq. (1), is written in terms of the Green's function, $\bar{g}(z|\xi; t)$, as

$$T_m^p(z, t) = \int_{\xi=0}^{z} \bar{g}_{z>\xi} F(\xi, t)d\xi + \int_{\xi=z}^{a_o(t)} \bar{g}_{z<\xi} F(\xi, t)d\xi \tag{9}$$

where

$$F(\xi, t) = -\frac{\rho L}{K_m}\{\psi_1 + \xi\psi_2\} \tag{10}$$

and the Green's function is

$$\bar{g}(z|\xi; t) = \begin{cases} \bar{g}_{z<\xi} = \xi - a_o \\ \bar{g}_{z>\xi} = z - a_o. \end{cases} \tag{11}$$

Substitution of Eq. (11) into Eq. (9) gives the temperature field corresponding to the particular solution.

The complete temperature field for the mush problem in Stage I in terms of the unknown functions $f_o(t)$ and $a_o(t)$ is thus

$$T_m(z, t) = T_l^* - \frac{Q_o a_o}{K_m}\left(1 - \frac{z}{a_o}\right) - \frac{\rho L}{2K_m}\left[\{z^2 - a_o^2\}\psi_1 + \frac{1}{3}\{z^3 - a_o^3\}\psi_2\right]. \tag{12}$$

Substitution of Eqs. (6) into Eq. (12) followed by substitution of the resulting expression into Eq. (7) gives

$$f_o(t) = \frac{1}{K_m(T_l^* - T_s^*)} \times \left[Q_o a_o - \frac{\rho L}{6}\{2a_o^2\dot{f}_o + a_o\dot{a}_o f_o\}\right]. \tag{13}$$

Application of Eq. (3b) gives

$$f_o(t) = \frac{2Q_o t}{\rho L a_o} \tag{14}$$

Substitution of Eq. (14) into Eq. (13) gives the mush region thickness

$$a_o(t) = \sqrt{\frac{4K_m(T_l^* - T_s^*)t}{\rho L}}. \tag{15}$$

Note that the initial condition, Eq. (2c), is satisfied automatically. Equations (14) and (15) give

$$f_o(t) = \sqrt{\frac{Q_o^2 t}{\rho L K_m(T_l^* - T_s^*)}}. \tag{16}$$

The final form of the thermal field in the mush region during Stage I results from substitution of Eqs. (15) and (16) into Eq. (12):

$$T_m(z, t) = T_l^* - \frac{Q_o(a_o - z)^2}{2a_o K_m}. \tag{17}$$

At time t_o, a continuous layer of solid has formed along the mold surface $z = 0$ and hence, $f_o(t_o) = 1$ and $T_m(0, t_o) = T_s^*$ (from Eq. (2b) which is the last condition to be satisfied). When these conditions are applied to Eq. (17), the following expression for t_o results:

$$t_o = \frac{\rho L K_m(T_l^* - T_s^*)}{Q_o^2}. \tag{18}$$

This completes the solution for the thermal field in the mush region during Stage I.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

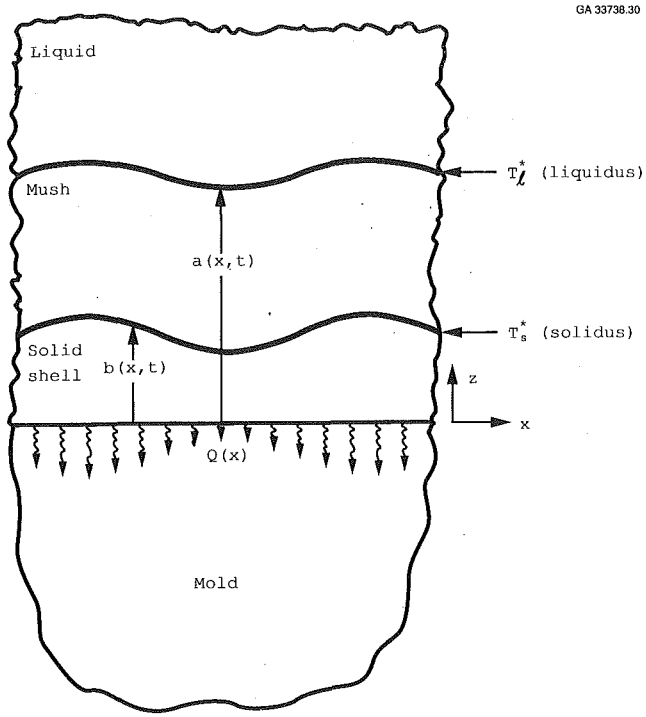


Fig. 2 Solidification system during Stage II

Stage II: Solid, Mush, and Liquid. Figure 2 shows the solidification system during Stage II where the solid shell is in contact with the mold surface and grows into the mush region. The cooling profile is no longer spatially uniform since we impose a small sinusoidal perturbation, $Q_1 \cos(mx)$, along the mold interface during Stage II. The resulting thermal field in the solid shell is given by $T_s(x, z, t)$. The mush/liquid and shell/mush boundaries lie at distances $a(x, t)$ and $b(x, t)$, respectively, above the mold surface. The thermal problem for the three region system in Fig. 2 is

$$\nabla^2 T_s(x, z, t) = 0; \quad t \geq t_o, \quad 0 < z < b(x, t) \tag{19}$$

$$\nabla^2 T_m(x, z, t) + \frac{\rho L}{K_m} \frac{\partial f(x, z, t)}{\partial t} = 0;$$

$$t \geq t_o, \; b(x, t) < z < a(x, t) \tag{20}$$

$$Q(x) = Q_o + Q_1 \cos(mx); \quad z = 0 \tag{21}$$

$$T_s(x, b, t) = T_s \tag{22}$$

$$K_s \frac{\partial T_s(x, b, t)}{\partial z} - K_m \frac{\partial T_m(x, b, t)}{\partial z} = \rho L(1 - f_{sb}) \frac{\partial b}{\partial t} \tag{23}$$

$$T_m(x, b, t) = T_s \tag{24}$$

$$T_m(x, a, t) = T_l^*; \quad \frac{\partial T_m(x, a, t)}{\partial z} = 0 \tag{25}$$

$$b(x, t_o) = 0; \quad a(x, t_o) = \frac{2K_m(T_l^* - T_s^*)}{Q_o}. \tag{26}$$

Note that Q_o and Q_1 are time-independent constants representing the uniform heat flux and the heat flux perturbation, respectively. This implies that the thermal problem is not coupled to the mechanical problem. As discussed by Li and Barber (1991), a drop in the contact pressure beneath thickness minima under these conditions implies that the coupled problem has positive feedback, but it is not a conclusive demonstration of the proposed instability mechanism. Nevertheless, we develop the uncoupled formulation since it points to a possible instability in the coupled problem. Equation (23) denotes a condition of nonequilibrium freezing where the shell/mush interface has a liquid fraction that freezes isothermally at the solidus temperature before the interface moves any further. The constant f_{sb} in Eq. (23) is the solid fraction at the shell/mush boundary. In most situations, f_{sb} is very close to one (see Tien and Geiger, 1967) and hence we simplify the remaining analysis by assuming that $f_{sb} = 1$, which denotes a condition of equilibrium freezing at the shell/mush interface. The constant m is the wave number and defines the wavelength ω of the perturbation through $\omega = 2\pi/m$. Shell thickness maxima lie above mold surface positions corresponding to $mx = \pm 2n\pi$ where n is an integer. The amplitudes of the perturbations in both fronts are assumed to be small compared to their wavelengths (i.e., $mb_1(t) \ll 1$, $ma_1(t) \ll 1$). Equation (23) therefore neglects the influence of freezing front curvature on the liberation of latent heat.

We anticipate the following forms for the shell and mush temperature fields, the shell/mush and mush/liquid interface positions relative to the mold surface, and the solid fraction distribution in the mush region:

$$T_s(x, z, t) = T_s^o(z, t) + T_s^1(z, t) \cos(mx) \tag{27}$$

$$T_m(x, z, t) = T_m^o(z, t) + T_m^1(z, t) \cos(mx) \tag{28}$$

$$b(x, t) = b_o(t) + b_1(t) \cos(mx) \tag{29}$$

$$a(x, t) = a_o(t) + a_1(t) \cos(mx) \tag{30}$$

$$f(x, z, t) = f_o(z, t) + f_1(z, t) \cos(mx) \tag{31}$$

where the first term on the right-hand side of each of Eqs. (27)–(31) corresponds to the zeroth-order solution due to uniform cooling while the remaining term corresponds to the first-order solution due to the spatial perturbation in the cooling profile. The zeroth-order and first-order problems for each region are determined through substitution of Eqs. (27)–(31) into Eqs. (19)–(26).

Zeroth-Order Thermal Field for the Solid Shell. The zeroth-order problem for the thermal field in the solid shell is

$$\frac{\partial^2 T_s^o(z, t)}{\partial z^2} = 0 \tag{32}$$

$$K_s \frac{\partial T_s^o(0, t)}{\partial z} = Q_o \tag{33}$$

$$T_s^o(b_o, t) = T_s^*; \quad K_s \frac{\partial T_s^o(b_o, t)}{\partial z} = K_m \frac{\partial T_m^o(b_o, t)}{\partial z} \tag{34}$$

$$b_o(t_o) = 0. \tag{35}$$

The solution to the above system of equations in terms of the unknown thickness $b_o(t)$ is

$$T_s^o(z, t) = T_s^* - \frac{Q_o b_o}{K_s}\left(1 - \frac{z}{b_o}\right) \tag{36}$$

where Eqs. (33) and (34a) have been used. The explicit form of $b_o(t)$ is to be determined from Eqs. (34b) and (35), which have yet to be applied. We will need the zeroth-order thermal field for the mush region, $T_m^o(z, t)$, before we can determine $b_o(t)$.

First-Order Thermal Field for the Solid Shell. The first-order problem for the thermal field in the solid shell is

$$\frac{\partial^2 T_s^1(z, t)}{\partial z^2} - m^2 T_s^1(z, t) = 0 \tag{37}$$

$$K_s \frac{\partial T_s^1(0, t)}{\partial z} = Q_1 \tag{38}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$T_s^1(b_o, t) = -b_1 \frac{\partial T_s^o(b_o, t)}{\partial z} \tag{39}$$

$$K_s\left[\frac{\partial T_s^1(b_o, t)}{\partial z} + b_1 \frac{\partial^2 T_s^o(b_o, t)}{\partial z^2}\right] = K_m\left[\frac{\partial T_m^1(b_o, t)}{\partial z} + b_1 \frac{\partial^2 T_m^o(b_o, t)}{\partial z^2}\right] \tag{40}$$

$$b_1(t_o) = 0. \tag{41}$$

Equations (39) and (40) were determined through a Taylor series expansion of Eqs. (22) and (23), respectively, about the mean solid shell thickness, $b_o(t)$. For more details on this procedure, the reader is referred to Li and Barber (1991). The solution to the above system of equations in terms of the unknown thicknesses $b_o(t)$ and $b_1(t)$ is

$$T_s^1(z, t) = \frac{Q_1}{mK_s}\sinh(mz) - \frac{\cosh(mz)}{\cosh(mb_o)}\left[\frac{Q_o}{K_s} b_1(t) + \frac{Q_1}{mK_s}\sinh(mb_o)\right] \tag{42}$$

where Eqs. (38) and (39) have been used. The explicit form of $b_1(t)$ is to be determined from Eq. (40) which has yet to be applied. We will need the zeroth-order and first-order thermal fields for both regions before we can determine the first-order solid shell thickness, $b_1(t)$.

Solid Fraction Distribution in Mush Region. The solid fraction distribution in the mush region during Stage II is assumed to be of the following form:

$$f(x, z, t) = \frac{a(x, t) - z}{a(x, t) - b(x, t)} \tag{43}$$

Substitution of Eqs. (29) and (30) into Eq. (43) gives Eq. (31), where

$$f_o(z, t) = \frac{a_o - z}{a_o - b_o};$$

$$f_1(z, t) = \frac{a_o b_1 - a_1 b_o}{[a_o - b_o]^2} + z\left\{\frac{a_1 - b_1}{[a_o - b_o]^2}\right\}. \tag{44}$$

Zeroth-Order Thermal Field in the Mush Region. The zeroth-order thermal problem for the mush region during Stage II is

$$\frac{\partial^2 T_m^o(z, t)}{\partial z^2} + \frac{\rho L}{K_m}\frac{\partial f_o(z, t)}{\partial t} = 0 \tag{45}$$

$$T_m^o(b_o, t) = T_s^* \tag{46}$$

$$T_m^o(a_o, t) = T_l^*; \quad \frac{\partial T_m^o(a_o, t)}{\partial z} = 0 \tag{47}$$

$$a_o(t_o) = \frac{2K_m(T_l^* - T_s^*)}{Q_o} \tag{48}$$

and Eq. (34*b*).

The thermal field corresponding to the homogeneous solution is determined through solution of the homogeneous form of Eq. (45) subject to Eqs. (46) and (47*a*). The particular solution for the zeroth-order thermal field is determined through solution of Eq. (45) with

$$\dot{f}_o(z, t) = \psi_3 + z\psi_4 \tag{49}$$

where

$$\psi_3(t) = \frac{a_o\dot{b}_o - \dot{a}_o b_o}{[a_o - b_o]^2}; \quad \psi_4(t) = \frac{\dot{a}_o - \dot{b}_o}{[a_o - b_o]^2} \tag{50}$$

and the homogeneous forms of Eqs. (46) and (47*a*). Using the Green's function method, the total zeroth-order temperature distribution is found to be

$$T_m^o(z, t) = \frac{T_s^* a_o - T_l^* b_o}{a_o - b_o} + z\left[\frac{T_l^* - T_s^*}{a_o - b_o}\right] - \frac{\rho L}{2K_m}\left[\psi_3\{z^2 - z[a_o + b_o] + a_o b_o\} + \frac{\psi_4}{3}\{z^3 - z[a_o^2 + a_o b_o + b_o^2] + a_o b_o[a_o + b_o]\}\right]. \tag{51}$$

Application of Eq. (47*b*) to Eq. (51) gives

$$a_o - b_o = \frac{6K_m}{\rho L}\left(\frac{T_l^* - T_s^*}{2\dot{a}_o + \dot{b}_o}\right). \tag{52}$$

Application of the remaining condition, Eq. (34*b*), gives

$$Q_o = K_m\left(\frac{T_l^* - T_s^*}{a_o - b_o}\right) + \frac{\rho L}{6}[\dot{a}_o + 2\dot{b}_o]. \tag{53}$$

Elimination of the quantity $a_o - b_o$ between Eqs. (52) and (53) gives

$$a_o + b_o = \frac{2Q_o t}{\rho L}. \tag{54}$$

Equation (54) is solved for $b_o(t)$ and the result is then substituted into Eq. (53) to generate the following differential equation for $a_o(t)$:

$$\left[a_o - \frac{Q_o t}{\rho L}\right]\dot{a}_o + \frac{2Q_o}{\rho L} a_o = \frac{2Q_o^2 t}{(\rho L)^2} + \frac{3K_m(T_l^* - T_s^*)}{\rho L}. \tag{55}$$

The solution to Eq. (55) is

$$a_o(t) = \frac{Q_o t}{\rho L} + \frac{K_m(T_l^* - T_s^*)}{Q_o}; \quad t \geq t_o \tag{56}$$

where Eq. (48) has been used as the initial condition. Equations (54) and (56) give

$$b_o(t) = \frac{Q_o t}{\rho L} - \frac{K_m(T_l^* - T_s^*)}{Q_o}; \quad t \geq t_o. \tag{57}$$

Using Eqs. (56) and (57), we define the nominal width of the mush region during Stage II, λ, as

$$\lambda = a_o - b_o = \frac{2K_m(T_l^* - T_s^*)}{Q_o}. \tag{58}$$

The time at which Stage II starts, t_o, is related to the width of the mush region during Stage II through

$$t_o(\lambda) = \frac{\rho L \lambda}{2Q_o}. \tag{59}$$

Substitution of Eqs. (50) into Eq. (51) gives the final form of the zeroth-order thermal field in the mush region during Stage II

$$T_m^o(z, t) = T_l^* - \frac{Q_o(a_o - z)^2}{2\lambda K_m}; \quad t \geq t_o. \tag{60}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

First-Order Thermal Field in the Mush Region. The thermal field corresponding to the first-order problem is determined by

$$\frac{\partial^2 T_m^1(z, t)}{\partial z^2} - m^2 T_m^1(z, t) + \frac{\rho L}{K_m}\frac{\partial f_1(z, t)}{\partial t} = 0 \quad (61)$$

$$T_m^1(b_o, t) = -b_1 \frac{\partial T_m^o(b_o, t)}{\partial z} \quad (62)$$

$$K_s\left[\frac{\partial T_s^1(b_o, t)}{\partial z} + b_1 \frac{\partial^2 T_s^o(b_o, t)}{\partial z^2}\right]$$

$$\left[\frac{m\lambda - \sinh(m\lambda)}{\sinh(m\lambda)}\right]\dot{b}_1 + \frac{mQ_o}{\rho L}\left[\frac{m^2\lambda^2 - \cosh(m\lambda) + 1}{\sinh(m\lambda)}\right]b_1$$

$$= \left[\frac{m\lambda\cosh(m\lambda) - \sinh(m\lambda)}{\sinh(m\lambda)}\right]\dot{a}_1 + \frac{mQ_o}{\rho L}\left[\frac{m\lambda\sinh(m\lambda) - \cosh(m\lambda) + 1}{\sinh(m\lambda)}\right]a_1. \quad (69)$$

Application of the remaining condition, Eq. (63), to Eq. (68) gives a second differential equation for the first-order thicknesses:

$$\left[\frac{m\lambda\cosh(m\lambda) - \sinh(m\lambda)}{\sinh(m\lambda)}\right]\dot{b}_1 + \frac{mQ_o}{\rho L}\left[\frac{m^2\lambda^2\cosh[m(\lambda + b_o)] + \cosh(mb_o)\cosh(m\lambda) - \cosh(mb_o)}{\sinh(m\lambda)\cosh(mb_o)} - m\lambda\right]b_1$$

$$= \left[\frac{m\lambda - \sinh(m\lambda)}{\sinh(m\lambda)}\right]\dot{a}_1 + \frac{mQ_o}{\rho L}\left[\frac{\cosh(m\lambda) - 1}{\sinh(m\lambda)}\right]a_1 + \frac{m^2\lambda^2 Q_1}{\cosh(mb_o)}. \quad (70)$$

$$= K_m\left[\frac{\partial T_m^1(b_o, t)}{\partial z} + b_1 \frac{\partial^2 T_m^o(b_o, t)}{\partial z^2}\right] \quad (63)$$

$$T_m^1(a_o, t) = -a_1 \frac{\partial T_m^o(a_o, t)}{\partial z};$$

$$\frac{\partial T_m^1(a_o, t)}{\partial z} = -a_1 \frac{\partial^2 T_m^o(a_o, t)}{\partial z^2} \quad (64)$$

$$a_1(t_o) = 0. \quad (65)$$

The thermal field corresponding to the homogeneous solution is determined through solution of the homogeneous form of Eq. (61) subject to Eqs. (62) and (64*a*). The thermal field corresponding to the particular solution is determined from Eq. (61) with

$$\dot{f}_1(z, t) = \psi_5 + z\psi_6 \quad (66)$$

and

$$\psi_5(t) = \frac{1}{\lambda^2}[\dot{a}_o b_1 + a_o\dot{b}_1 - \dot{a}_1 b_o - a_1\dot{b}_o];$$

$$\psi_6(t) = \frac{1}{\lambda^2}[\dot{a}_1 - \dot{b}_1]. \quad (67)$$

Using the Green's function method, the total first-order thermal field in the mush region during Stage II is found to be

$$T_m^1(z, t) = \frac{Q_o b_1}{K_m}\frac{\sinh[m(z - a_o)]}{\sinh(m\lambda)} - \frac{\rho L}{m^2 K_m}\left[\begin{array}{l}\psi_5(t)\left\{\dfrac{\sinh[m(z - b_o)] - \sinh[m(z - a_o)]}{\sinh(m\lambda)} - 1\right\} \\ + \psi_6(t)\left\{\dfrac{a_o\sinh[m(z - b_o)] - b_o\sinh[m(z - a_o)]}{\sinh(m\lambda)} - z\right\}\end{array}\right] \quad (68)$$

where it can be shown that $T_m^1(z, t_o) = 0$. Application of Eq. (64*b*) to Eq. (68) gives the following differential equation for the first-order thicknesses $a_1(t)$ and $b_1(t)$:

The pure metal limit $\lambda \to 0$ is excluded from the present model since the condition $b_1(t) > \lambda$ occurs in the limiting process. The origin of this problem lies with the fact that we have expanded the boundary conditions, Eqs. (22)–(25), about the mean fronts $a_o(t)$ and $b_o(t)$. A cursory examination of Eqs. (56)–(60) reveals that the zeroth-order thermal field for the alloy problem reduces to its pure metal counterpart as $\lambda \to 0$. Hence, the pure metal limit is excluded from the model through the first-order analysis.

The Thermal Stress Problem

Since the applied heat flux Eq. (21) contains a spatially periodic term, we assume the thermal stress distribution in the solid shell, σ_{ij}, and the contact pressure at the shell/mold interface, P, adopt the following forms:

$$\sigma_{ij}(x, z, t) = \sigma_{ij}^o(z, t) + \sigma_{ij}^1(z, t)\cos(mx) \quad (71)$$

$P(x, t) = P_o(t) + P_1(t)\cos(mx)$ and

$$P_o(t) = -\sigma_{zz}^o(0, t);\ P_1(t) = -\sigma_{zz}^1(0, t). \quad (72)$$

The goals of the thermal stress problem are first to calculate the contact pressure P beneath shell thickness minima for selected values of the nominal freezing range λ and then determine whether or not $P = 0$.

Stress accumulation in the solid shell is assumed to begin with the start of Stage II. However, there is a region within the mush that lies between the solidus temperature and the coherency temperature wherein stress accumulation begins due to lateral interconnection of adjacent dendrites (Ackermann et al., 1985). We neglect this physics in order to illustrate salient aspects of the freezing range effect on the contact pressure at the mold surface.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

As discussed by Li and Barber (1991), the total stress distribution in the solid shell can be expressed as a linear combination of a particular solution, σ_{ij}^p, that corresponds to the thermal field, an isothermal solution, σ_{ij}^h, which is allowed to vary in time so as to satisfy time-varying terms in the boundary conditions, and a residual stress, σ_{ij}^r, which is the stress that remains in the solid shell after it is cooled to a uniform temperature and relieved of all boundary tractions. The stress field corresponding to the particular solution is derived from the thermoelastic displacement potential, φ, through

$$\sigma_{xx}^p = -\frac{\partial^2\varphi}{\partial z^2}; \quad \sigma_{zz}^p = -\frac{\partial^2\varphi}{\partial x^2}; \quad \sigma_{xz}^p = \frac{\partial^2\varphi}{\partial x\partial z}. \tag{73}$$

The corresponding equilibrium relation is

$$\nabla^2\varphi = 2E\alpha T_s \tag{74}$$

where we assume that deformation of the solid shell occurs at constant volume and hence $\nu = 0.5$ for the remainder of the analysis. The displacement field, $\vec{u}^p$, corresponding to the particular solution is related to φ through

$$\vec{\nabla}\varphi = \frac{2E\vec{u}^p}{3}. \tag{75}$$

The isothermal and residual stress fields may be derived from the Airy stress function, Φ, according to

$$\sigma_{xx}^h = \frac{\partial^2\Phi}{\partial z^2}; \quad \sigma_{zz}^h = \frac{\partial^2\Phi}{\partial x^2}; \quad \sigma_{xz}^h = -\frac{\partial^2\Phi}{\partial x\partial z}. \tag{76}$$

The Airy stress function must satisfy the following compatibility relation:

$$\frac{\partial}{\partial t}\nabla^4\Phi = 0. \tag{77}$$

The mechanical boundary conditions for frictionless contact at the mold surface are

$$\sigma_{zz} = 0; \quad \dot{u}_z = 0; \quad z = 0. \tag{78}$$

Solidification at the shell/mush interface is assumed to occur in a state of hydrostatic stress. Hence,

$$\sigma_{xx} = \sigma_{zz} = -p, \quad \sigma_{xz} = 0; \quad z = b_o(t) \tag{79}$$

where p is the fluid pressure at the shell/mush interface.

Zeroth-Order Thermal Stress Field. The zeroth-order thermal stress field is

$$\sigma_{xx}^o = -p + \frac{2E\alpha Q_o b_o}{K_s}\left(1 - \frac{z}{b_o}\right);$$

$$\sigma_{zz}^o = -p; \quad \sigma_{xz}^o = 0. \tag{80}$$

First-Order Thermal Stress Field. The particular solution to the first-order thermal stress field is derived from

$$\nabla^2\varphi = 2E\alpha T_s^1 \cos(mx) \tag{81}$$

where T_s^1 is given by Eq. (42). A particular solution of Eq. (81) is

$$\varphi(x, z, t) = [D_1 \cosh(mz) + D_2(t) \sinh(mz)]z \cos(mx) \tag{82}$$

where

$$D_1 = \frac{E\alpha Q_1}{m^2K_s};$$

$$D_2(t) = -\frac{E\alpha}{m^2K_s \cosh(mb_o)}[Q_o mb_1 + Q_1 \sinh(mb_o)]. \tag{83}$$

The stress field corresponding to the particular solution is thus determined from Eqs. (73) and (82).

From Eqs. (75), the normal velocity in terms of the displacement potential is

$$\dot{u}_{z_1}^p = \frac{3}{2E}\frac{\partial}{\partial t}\left(\frac{\partial\varphi}{\partial z}\right). \tag{84}$$

The isothermal stress field is derived from the following Airy stress function Φ:

$$\Phi(x, z, t) = \{[C_1(t) + zC_2(t)] \cosh(mz) + [C_3(t) + zC_4(t)] \sinh(mz) + g(z)\} \cos(mx) \tag{85}$$

where the $C_i(t)$ are unknown functions to be determined in the analysis and $g(z)$ is a function that represents residual stress. The isothermal stress field is thus determined in terms of the unknown $C_i(t)$ with Eqs. (76) and (85).

The normal velocity corresponding to the isothermal solution is determined via the plane-strain constitutive relation for elastic materials

$$\dot{u}_{z_1}^h = \frac{1+\nu}{E}\left[(1-\nu)\int \dot{\sigma}_{zz_1}^h dz - \nu\int \dot{\sigma}_{xx_1}^h dz\right] \tag{86}$$

where we will set $\nu = 0.5$.

Application of Eq. (78a) to the total shear stress (the sum of Eqs. (73c) and (76c)) and Eq. (78b) to the total normal velocity (the sum of Eqs. (84) and (86)) gives

$$C_2(t) = 0; \quad C_3(t) = \frac{E\alpha Q_1}{m^3K_s}. \tag{87}$$

Application of the hydrostatic boundary conditions Eqs. (79) to the first-order problem requires that we first expand the stress field at the shell/mush interface about $b_o(t)$ in a Taylor series. Since the details of this procedure are in the literature, we present only the results from this process. The pertinent boundary conditions on the first-order stress field at $b_o(t)$ are

$$(\sigma_{xx_1}^p + \sigma_{zz_1}^h)|_{z=b_o} = -b_1\left(\frac{\partial\sigma_{xx}^o}{\partial z}\right)_{z=b_o}\cos(mx);$$

$$(\sigma_{xz_1}^p + \sigma_{xz_1}^h)|_{z=b_o} = 0; \quad (\sigma_{zz_1}^p + \sigma_{zz_1}^h)|_{z=b_o} = 0. \tag{88}$$

Substitution of the total stress components (the explicit forms of which are omitted due to their length) into Eqs. (88) gives the remaining coefficients $C_1(t)$ and $C_4(t)$ and

$$\frac{1}{2m^2}[mb_o w^2 - mb_o - w]g''(b_o) + \frac{1}{m}g'(b_o) - \tfrac{1}{2}[mb_o w^2 - mb_o + w]g(b_o) = \frac{E\alpha Q_o}{m^2K_s}[mb_o b_1 w^2 - mb_o b_1 - wb_1] - \frac{E\alpha w b_o Q_1}{m^2K_s}\sinh(mb_o) \tag{89}$$

where $w = \tanh(mb_o)$, $A = 2E\alpha/K_s$. Equation (89) is a differential equation for the residual stress function $g(z)$.

Solution for the First-Order Contact Pressure. For convenience in the numerical solution for $P_1(t)$, we reformulate the pertinent equations in terms of the following dimensionless quantities:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\eta = mz; \quad \bar{a}_1 = ma_1; \quad \bar{b}_1 = mb_1; \quad \bar{\lambda} = m\lambda;$$

$$\bar{Q}_{10} = Q_1/Q_o; \quad G(\eta) = \frac{m^3 K_s}{2E\alpha Q_1} g(z)$$

$$\bar{C}_1(\eta) = \frac{m^3 K_s}{E\alpha Q_1} C_1(t); \quad \bar{P}_1(\eta) = \frac{mK_s}{2E\alpha Q_1} P_1(t).$$

The dimensionless form of Eq. (89) is

$$\tfrac{1}{2}[\eta w^2 - \eta - w]G''(\eta) + G'(\eta) - \tfrac{1}{2}[\eta w^2 - \eta + w]G(\eta)$$

$$= -\frac{\bar{Q}_{10}^{-1}}{2}[(1 - w^2)\eta + w]\bar{b}_1 - \frac{w\eta}{2}\,\text{sech}\,(\eta). \quad (90)$$

Note that Eq. (89) has been recast in the z-domain, where $z = b_o(t)$, prior to being normalized. The dimensionless form of $C_1(t)$ is

$$\bar{C}_1(\eta) = \frac{1}{\cosh\,(\eta)}[\eta w\{G''(\eta) - G(\eta)\}$$

$$- 2G(\eta) - \bar{Q}_{10}^{-1} w\eta\bar{b}_1 + \eta\,\text{sech}\,(\eta) - \sinh\,(\eta)]. \quad (91)$$

The dimensionless form of the first-order contact pressure is

$$\bar{P}_1(\eta) = \bar{C}_1(\eta)/2 + G(0). \quad (92)$$

The solution procedure for $\bar{P}_1(\eta)$ begins with numerical solution of the dimensionless forms of Eqs. (69) and (70) for $\bar{a}_1$ and $\bar{b}_1$ with a selected value of $\bar{\lambda}$. The calculated $\bar{b}_1$ is used in Eq. (90) which is solved for the residual stress function $G(\eta)$ according to the initial conditions $G(0) = 0$ and $G'(0) = 0$. The dimensionless Airy stress function coefficient, $\bar{C}_1(\eta)$, which is evaluated with the calculated $G(\eta)$, is then used in Eq. (92) to calculate $\bar{P}_1(\eta)$. Since Eq. (92) gives the first-order contact pressure beneath thickness maxima, we multiply it by -1 and then add the result to a specified $\bar{P}_o$ to study the contact pressure beneath thickness minima.

Results and Concluding Remarks

We examine the evolution of the contact pressure beneath thickness minima at selected values of the zeroth-order pressure, $\bar{P}_o$. Air gap nucleation occurs when the contact pressure falls to zero thereby signaling the onset of unstable growth of the freezing front. After this time, the analysis is terminated. From Eq. (72a), air gap nucleation occurs when

$$\bar{P}_o = \bar{P}_1. \quad (93)$$

For each case considered, $\bar{Q}_{10} = 10^{-5}$ (although $\bar{Q}_{10}$ has little impact on the instability problem). The Runge-Kutta technique was used to generate the numerical solutions.

Figure 3 shows the evolution of the contact pressure, $\bar{P}$, for selected values of the nominal freezing range width, $\bar{\lambda}$, and $\bar{P}_o = 0.05$. For the sake of comparison, the curve corresponding to $\bar{\lambda} = 0$ is included even though the first-order part of the present theory does not include the pure metal case. This curve corresponds to the solutions developed by Richmond et al. (1990) and Li and Barber (1991). The contact pressure falls to zero in all cases signaling air gap nucleation beneath thickness minima, but the time at which air gap nucleation occurs is dependent upon $\bar{\lambda}$. Air gap nucleation occurs sooner for those alloys with the larger freezing ranges. Hence, unstable growth of the pure metal takes the longest time to manifest itself for the liquid pressure considered in Fig. 3.

Figure 4 shows the evolution of the contact pressure, $\bar{P}$, for selected values of the nominal freezing range width, $\bar{\lambda}$, and $\bar{P}_o = 0.4$. The trend displayed in Fig. 3 has reversed itself as the pure metal will nucleate air gaps before any of the alloys. As $\bar{\lambda}$ increases, the time to air gap nucleation increases. For $\bar{\lambda} = 2.0$, 10.0, 15.0, the contact pressure levels off to a constant value with increasing time once the thickness of the solid shell exceeds the wavelength of the spatial perturbation in the heat flux profile. At this point in the process, the contact pressure perturbation has little influence on the macromorphology of the solid shell: air gap nucleation never occurs for these conditions. Also, for alloys where $\bar{\lambda} > 10$, the average value of the contact pressure at longer times approaches a constant value (e.g., the $\bar{\lambda} = 10.0$, 15.0 curves are very close) and hence shell growth instability is no longer sensitive to the nominal freezing range.

Figures 3 and 4 imply that $\bar{P}_o$ plays a prominent role in the air gap nucleation process for the range of $\bar{\lambda}$ considered. The conditions described in Fig. 3 are in qualitative agreement with the experiments of Singh and Blazek (1974) since the wider freezing range alloys are more susceptible to growth instability

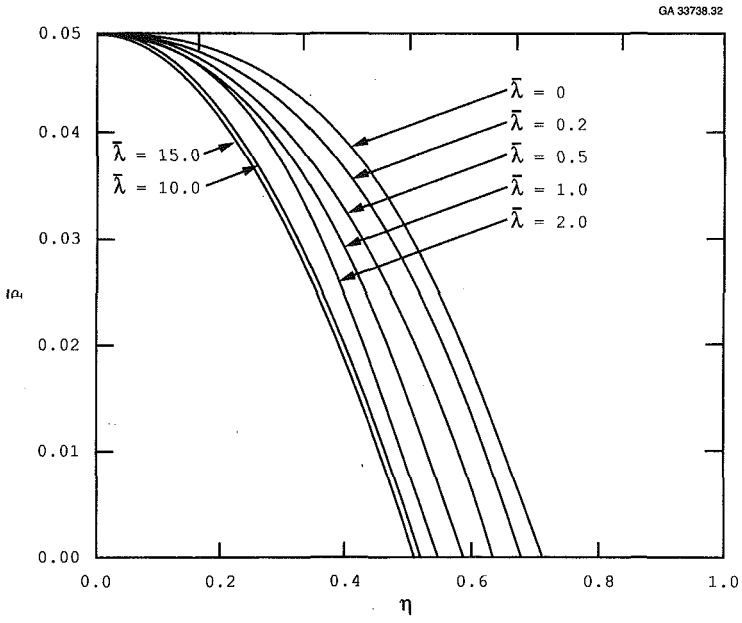


Fig. 3 Evolution of contact pressure beneath a thickness minimum for $\bar{P}_o(\eta) = 0.05$ at selected values of $\bar{\lambda}$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

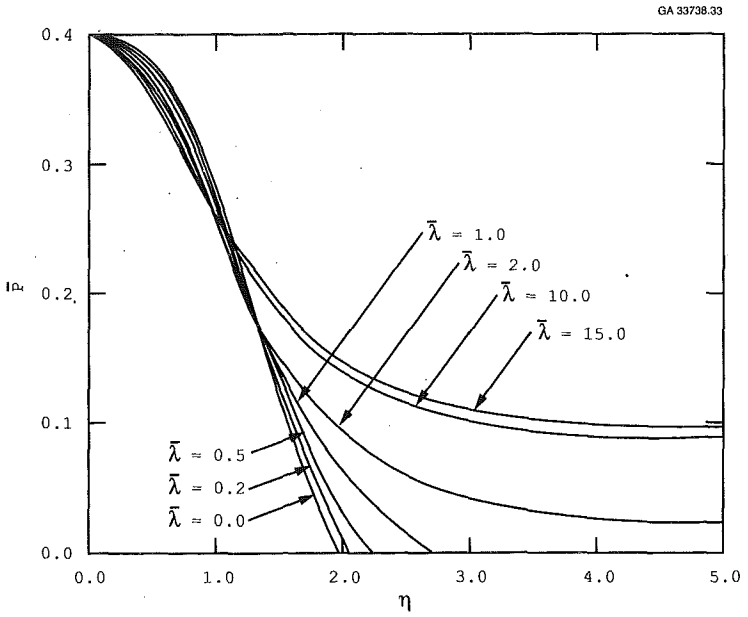


Fig. 4 Evolution of contact pressure beneath a thickness minimum for $\bar{P}_o(\eta) = 0.4$ at selected values of $\bar{\lambda}$

than the narrower freezing range alloys. However, when the zeroth-order contact pressure is increased by nearly an order of magnitude, the trend in Fig. 3 reverses itself as shown in Fig. 4. Those alloys with the narrowest freezing ranges tend to be more susceptible to growth instability than the wider freezing range alloys, a result that is in qualitative agreement with the experiments of Cisse et al. (1971). Since the experimental results of Singh and Blazek and Cisse et al. did not include effects of molten metal pressure, further experimentation will be necessary to come to a definitive conclusion about the freezing range effect on alloy growth instability and the role that the pressure plays.

As mentioned previously, the first-order part of the present theory does not include the pure metal case. It would be of immediate interest to re-pose the thermal and mechanical problems in the curvilinear coordinate systems that match the spatially oscillatory fronts $a_1(t)$ and $b_1(t)$. The mathematical complexity associated with this approach will be substantially greater than that associated with the present work. Moreover, it is anticipated that the major results from such a model will be virtually identical to those presented herein. The physics of the model could be improved through inclusion of the associated solute diffusion problem although this is not expected to effect the major results of the model.

References

Ackermann, P., Kurz, W., and Heinemann, W., 1985, "In Situ Tensile Testing of Solidifying Aluminum and Al-Mg Shells," *Matrls. Sci. Eng.*, Vol. 75, pp. 79–86.

Bäckerud, L., Krol, E., and Tamminen, J., 1986, *Solidification Characteristics of Aluminum Alloys, Vol. 1: Wrought Alloys.* SkanAluminum, Oslo, Norway.

Cisse, J., Cole, G. S., and G. F. Bolling, 1971, "Freezing Front Asymmetry During Ingot Solidification of Aluminum and its Alloys," *AFS Cast Metals Res. J.*, Vol. 7, pp. 158–161.

DeGarmo, E. P., Black, J. T., and Kohser, R. E., 1988, *Materials and Processes in Manufacturing*, 7th ed., MacMillan, New York.

Li, N.-Y., and Barber, J. R., 1991, "Thermoelastic Instability in Planar Solidification," *Internat. J. Mech. Sci.*, Vol. 33, pp. 945–959.

Richmond, O., Hector, Jr., L. G., and Fridy, J. M., 1990, "Growth Instability During Nonuniform Directional Solidification of Pure Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp. 529–536.

Stakgold, I., 1967, *Boundary Value Problems of Mathematical Physics.* Vol. 1, MacMillan, New York.

Singh, S. N., and Blazek, K. E., 1974, "Heat Transfer and Skin Formation in a Continuous-Casting Mold as a Function of Steel Carbon Content," *J. of Metals*, Vol. 26, pp. 17–27.

Tien, R. H., and Geiger, G. E., 1967, "A Heat Transfer Analysis of the Solidification of a Binary Eutectic System," ASME *Journal of Heat Transfer*, Vol. 89, pp. 230–234.

Yigit, F., and Barber, J. R., 1992, "Effect of Thermal Capacity on the Development of Sinusoidal Perturbation During the Unidirectional Solidification of Pure Metals," ASME HTD Vol. 218/AMD-Vol. 139, pp. 141–150.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stability of Generalized Equilibria in Constrained Pseudodissipative Systems

J. A. Walker
Department of Mechanical Engineering,
Northwestern University,
Evanston, IL 60208

We present improved stability criteria for the generalized equilibria of pseudodissipative systems. Compared to earlier results (Walker, 1988), the new criteria apply to many more kinematically constrained systems. We also obtain particularly simple stability criteria for holonomically constrained systems. Minimality is not assumed and linearization is not employed.

1 Introduction

Stability criteria for the generalized equilibria of constrained mechanical systems were presented in Theorems 3.1–3.2 of (Walker, 1988), where emphasis was placed upon the simplest possible stability criteria, presented in the form of corollories restricted to locally unconstrained equilibria. For constrained equilibria, even the general stability criteria in Walker (1988) may be too restrictive to be useful. This work is intended to rectify that problem.

By using improved Liapunov functions in Section 3, we obtain general stability criteria applicable to a much wider class of constrained systems. Compared to the theorems in Walker (1988), our new Theorems 3.1–3.2 place very few restrictions on the form of the kinematic constraints. In Section 4 we obtain simplified corollaries of these theorems. Corollaries 4.1–4.2 are applicable to almost all constrained systems, unlike the "unconstrained" corollaries in Walker (1988). The even simpler Corollaries 4.3–4.4 are limited to holonomically constrained systems. In Section 5 the usefulness of these results is demonstrated by means of an example for which the constraints seem "unremovable."

In the following Section 2 we first discuss a widely accepted kinematic claim which seems to be untrue and, even if true, is impractical for many mechanisms. This discussion explains our interest in obtaining simple stability criteria for holonomically constrained systems. We then provide a general constrained formulation of Lagrangian dynamics for pseudodissipative systems. This was also done in Walker (1988, 1989), but here we relax the definition of "pseudodissipative" somewhat, making it more convenient to use with constrained systems.

2 Constrained Pseudodissipative Systems

It is often claimed that for any mechanism there exists a set of generalized coordinates whose number is "minimal," in the sense that there remain no holonomic constraints on these coordinates. See Likens (1973), Rosenberg (1977), and Kane and Levinson (1985). Although a few authors have made serious efforts to justify this claim (Rosenberg, 1977), their arguments are local and fail at some locations. The general claim seems to have no general proof, and it may be untrue.

Despite many attempts, the author has been unable to devise any one coordinate which uniquely determines the position (θ, ψ) of the slider/crank mechanism shown in Fig. 1; the holonomic constraint on (θ, ψ) seems not to be "removable." The author's failure does not prove that success is impossible, but it demonstrates the impracticality of requiring minimality. The minimality assumption severely limits the usefulness of many treatises on the subject of Lagrangian dynamics, which we now formulate without this assumption.

Consider a collection of particles observed by some inertial observer, and choose some n "generalized coordinates" $(q_1, q_2, \ldots, q_n) \equiv q \in \mathcal{O} \subset \mathcal{R}^n$, where $\mathcal{O}$ is some open subset of $\mathcal{R}^n$; the dimension n of the *generalized position* $q \in \mathcal{R}^n$ need *not* be minimal. We denote by $u \in \mathcal{R}^n$ the corresponding *generalized speed;* i.e., $u(t) = \dot{q}(t)$ along motions $q(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$. The resulting *generalized kinetic energy* T: $\mathcal{R} \times \mathcal{O} \times \mathcal{R}^n \to \mathcal{R}$ and *generalized force* Q: $\mathcal{R} \times \mathcal{O} \times \mathcal{R}^n \to \mathcal{R}^n$ depend upon the *generalized state* $(q, u) \in \mathcal{O} \times \mathcal{R}^n$ and possibly the current time $t \in \mathcal{R}$. Any and all kinematic constraints on (q, u) may be accommodated by defining a *kinematically possible set* $\mathcal{C}(t) \subset \mathcal{O} \times \mathcal{R}^n$, consisting of all generalized states (q, u) kinematically possible at time $t \in \mathcal{R}$. A C^1-smooth function $q(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$ is *kinematically possible on* $[t_1, t_2)$ if $(q(t), \dot{q}(t)) \in \mathcal{C}(t)$ for all $t \in [t_1, t_2)$.

If $\mathcal{C}(t)$ is a strict subset of $\mathcal{O} \times \mathcal{R}^n$, $t \in \mathcal{R}$, it is so described by a number of kinematic constraints. Each constraint on q implies another constraint on u, since $u(t) \equiv \dot{q}(t)$ along motions; such a constraint is called "holonomic." There also may be "nonholomonic" constraints on u, not following from corresponding constraints on q. For example, a nonholomonic constraint is created when we assume a disk to roll on some given surface, unless the rolling contact path is known beforehand.

For any constrained system, the motion equations of J. L. Lagrange (1788) may be obtained by replacing the pair (L, D) by (T, Q) in the subsequent Theorem 2.1. As written, Theorem 2.1 is intended for a special class of systems called "pseudodissipative" (Walker, 1988). In the following definition, and whenever notationally convenient thereafter, real n-tuples (elements of $\mathcal{R}^n$) are identified with column matrices.

Definition 2.1: The generalized force Q is called *pseudodissipative* if there exist a C^1-smooth *pseudopotential* U: $\mathcal{R} \times \mathcal{O} \times \mathcal{R}^n \to \mathcal{R}$, affine with respect to its third argument, and a *dissipative part* D: $\mathcal{R} \times \mathcal{O} \times \mathcal{R}^n \to \mathcal{R}^n$, such that the following conditions hold:

(a) For all $(q, u) \in \mathcal{C}(t)$, $(q, v) \in \mathcal{C}(t)$, $t \in \mathcal{R}$,

$$[u - v]^T[D(t, q, u) - D(t, q, v)] \leq 0. \tag{1}$$

(b) Along every kinematically possible $q(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$,

$$[\delta(t)]^T\left[\frac{d}{dt}\frac{\partial}{\partial u}U(t, q(t), \dot{q}(t)) - \frac{\partial}{\partial q}U(t, q(t), \dot{q}(t))\right]$$

$$= [\delta(t)]^T[Q(t, q(t), \dot{q}(t)) - D(t, q(t), \dot{q}(t))] \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Feb. 15, 1995; final revision, Oct. 15, 1995. Associate Technical Editor: S. W. Shaw.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

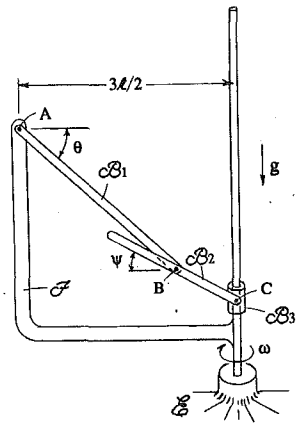


Fig. 1 Mechanism of example in Section 5

for all C^1-smooth $\delta(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$ such that $(q(t), \dot{q}(t) + \delta(t)) \in \mathcal{C}(t)$ for all $t \in \mathcal{R}$.

The corresponding *Lagrangian* is $L \equiv T - U$, and Q is *pseudoconservative* if (b) holds with $D(t, q, u) \equiv 0$. Q is *strongly pseudodissipative* if equality occurs in (1) *only* for $u = v$ at each $t \in \mathcal{R}$.

For constrained systems the conditions of Definition 2.1 are slightly weaker than those of the original definition in (Walker, 1988, 1989, 1990). It can be shown that the generalized force is pseudodissipative if and only if Q meets the monotonicity condition (a) required of D; we *could* choose $U \equiv 0$ and $D \equiv Q$, since the pair (U, D) of Definition 2.1 is far from unique. Definition 2.1 is stated as it is because the idea of pseudodissipative Q is most useful if D depends little or not at all on its second argument $q \in \mathcal{Q}$, and this often requires nontrivial U to be chosen.

For pseudodissipative Q: $\mathcal{R} \times \mathcal{Q} \times \mathcal{R}^n \to \mathcal{R}^n$, Lagrange's formulation produces the following consequence of Newton's second law:

Theorem 2.1: If a C^1-smooth function $q(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$ is dynamically possible on $[t_o, t_f)$, then for all $t \in [t_o, t_f)$,

$$(q(t), u(t)) \in \mathcal{C}(t), \; u(t) \equiv \dot{q}(t), \tag{3}$$

and

$$0 = [\delta(t)]^T \left[\frac{d^+}{dt} \frac{\partial}{\partial u} L(t, q(t), u(t)) - \frac{\partial}{\partial q} L(t, q(t), u(t)) - D(t, q(t)u(t)) \right] \tag{4}$$

for all C^1-smooth $\delta(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$ such that $(q(t), u(t) + \delta(t)) \in \mathcal{C}(t)$, $t_o \leq t < t_f$.

Henceforth we shall assume that

(i) a particular initial instant t_o has been chosen and the time-interval of interest is $[t_o, \infty)$,
(ii) Q is known and pseudodissipative on $[t_o, \infty)$,
(iii) for each $(q^o, u^o) \in \mathcal{C}(t_o)$, there exists a unique C^1-smooth solution $q(\cdot)$: $[t_o, \infty) \to \mathcal{R}^n$ of (3)–(4) such that $q(t_o) = q^o$ and $u(t_o) = u^o$. Moreover, at each $t > t_o$, $(q(t), u(t))$ depends continuously on (q^o, u^o).

Obviously, the simplest type of solution of (3)–(4) would be a constant solution, $q(t) \equiv q^e$ for some fixed $q^e \in \mathcal{R}^n$; the corresponding initial generalized state $(q^e, 0) \in \mathcal{C}(t_o)$ is called a *generalized equilibrium.*

Corollary 2.1: If L is time invariant, then $(q^e, 0)$ is a generalized equilibrium if and only if, for all $t \geq t_o$, $(q^e, 0) \in \mathcal{C}(t)$ and

$$0 = [\delta(t)]^T \left[\frac{\partial}{\partial q} L(t_o, q^e, 0) + D(t, q^e, 0) \right] \tag{5}$$

for every C^1-smooth $\delta(\cdot)$: $\mathcal{R} \to \mathcal{R}^n$ such that $(q^e, \delta(t)) \in \mathcal{C}(t)$ for $t \geq t_o$.

3 General Stability Criteria

In all of the following stability results we shall assume that Q is pseudodissipative with time-invariant L and D. At each generalized equilibrium $(q^e, 0) \in \mathcal{R}^n \times \mathcal{R}^n$, we shall employ three $(q^e, 0)$-related functions defined on $\mathcal{Q} \times \mathcal{R}^n$:

$$G(q, u) \equiv -L(t_o, q, u) + u^T \frac{\partial}{\partial u} L(t_o, q, u) - q^T D(t_o, q^e, 0), \tag{6}$$

$$J(q, u; \gamma) \equiv [\gamma(q)]^T \left[\frac{\partial}{\partial u} L(t_o, q, u) - \frac{\partial}{\partial u} L(t_o, q^e, 0) \right], \tag{7}$$

$$R(q, u; \gamma) \equiv \sum_{i=1}^{n} u_i \left[\frac{\partial}{\partial q_i} \gamma(q) \right]^T \times \left[\frac{\partial}{\partial u} L(t_o, q, u) - \frac{\partial}{\partial u} L(t_o, q^e, 0) \right] + [\gamma(q)]^T \left[\frac{\partial}{\partial q} L(t_o, q, u) + D(t_o, q, u) \right], \tag{8}$$

where γ: $\mathcal{Q} \to \mathcal{R}^n$ is any locally C^1-smooth function. We shall also employ two $(q^e, 0)$-related open sets parameterized by $h > 0$,

$$\mathcal{N}_h \equiv \{(q, u) \in \mathcal{R}^n \times \mathcal{R}^n | 0 < |q - q^e|^2 + |u|^2 < h\}, \tag{9}$$

$$\Omega_h \equiv \{(q, u) \in \mathcal{N}_h | G(q, u) < G(q^e, 0)\}. \tag{10}$$

For constrained systems the functions G, J, and R differ from those defined in Walker (1988), and so do the following two theorems. Our new Theorems 3.1–2 follow from the motion equations of Theorem 2.1 and the same Liapunov-type arguments used in Walker (1988).

Theorem 3.1: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$:

(a) $\mathcal{N}_h \cap \mathcal{C}(t)$ is time invariant, and the statement $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$ implies that $(q, 0) \in \mathcal{C}(t_o)$.
(b) L and D are time invariant, and $u^T[D(t_o, q, 0) - D(t_o, q^e, 0)] = 0$ for all $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$.
(c) $G(q, u) > G(q^e, 0)$ (strictly) for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$, where $\bar{\mathcal{C}}(t_o)$ denotes the closure of $\mathcal{C}(t_o)$.

Then $(q^e, 0)$ is *stable.* If Q is pseudoconservative, then $(q^e, 0)$ is *not* asymptotically stable. Suppose that the following condition also holds:

(d) There does *not* exist a solution of Eqs. (3)–(4) along which $(q(t), u(t)) \in \mathcal{N}_h$, $G(q(t), u(t))$ is constant, and $[u(t)]^T[D(t_o, q(t), u(t)) - D(t_o, q^e, 0)] = 0$ for all $t \geq t_o$.

Then $(q^e, 0)$ is *asymptotically* stable.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

If the foregoing conditions (a) and (b) hold, (d) holds with $\mathcal{N}_h$ replaced by Ω_h, but in place of (c) we have its strong contradiction,

(e) The set $\Omega_\epsilon \cap \mathcal{C}(t_o)$ is nonempty for every $\epsilon > 0$, then $(q^e, 0)$ is *unstable.*

Theorem 3.2: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$ and some locally C^1-smooth function γ: $\mathcal{D} \to \mathcal{R}^n$:

(a) $\Omega_h \cap \mathcal{C}(t)$ is time invariant and the statement $(q, u) \in \Omega_h \cap \mathcal{C}(t_o)$ implies that both $(q, 0) \in \mathcal{C}(t_o)$, $(q, u + \gamma(q)) \in \mathcal{C}(t)$ for every $t \geq t_o$.

(b) L and D are time invariant, and $u^T[D(t_o, q, 0) - D(t_o, q^e, 0)] = 0$ for all $(q, u) \in \Omega_h \cap \mathcal{C}(t_o)$.

(c) The set $\Omega_\epsilon \cap \mathcal{C}(t_o)$ is nonempty for every $\epsilon > 0$.

(d) There exists a real number $\mu \geq 0$ such that $R(q, u; \gamma) - \mu u^T[D(t_o, q, u) - D(t_o, q, 0)] \geq 0$ for all $(q, u) \in \Omega_h \cap \mathcal{C}(t_o)$.

(e) There does *not* exist a solution of Eqs. (3)–(4) along which $(q(t), u(t)) \in \Omega_h$, $G(q(t), u(t))$ and $J(q(t), u(t); \gamma)$ are constant, and $R(q, (t), u(t); \gamma) = 0 = [u(t)]^T[D(t_o, q(t), u(t)) - D(t_o, q^e, 0)]$ for all $t \geq t_o$.

Then $(q^e, 0)$ is *unstable.*

Comparing each of Theorems 3.1–3.2 with its correspondent in (Walker, 1988), we see that conditions (a)–(b) have been significantly weakened for constrained systems. This improvement is due to the improved Liapunov functions G and $J + \mu G$ used in Theorems 3.1–3.2.

4 Simpler Stability Criteria

Theorem 3.1 and 3.2 are our most general stability criteria, but they can be very difficult to apply. Simplified results, more convenient in applications, are presented in the following Corollaries 4.1–4.4. However, we first need to provide some notation.

We define $(q^e, 0)$-related $n \times n$ matrices $M = [m_{ij}]$, $K = [k_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$, $\bar{C}$, and X as follows:

$$m_{ij} \equiv \frac{\partial^2}{\partial u_i \, \partial u_j} L(t_o, q^e, 0),$$

$$k_{ij} \equiv -\frac{\partial^2}{\partial q_i \, \partial q_j} L(t_o, q^e, 0), \tag{11}$$

$$b_{ij} \equiv \frac{\partial^2}{\partial q_i \, \partial u_j} L(t_o, q^e, 0), \quad c_{ij} \equiv -\frac{\partial}{\partial u_j} D_i(t_o, q^e, 0), \tag{12}$$

$$2\bar{C} \equiv C + C^T, \; 4X \equiv (2B - C) - (2B - C)^T. \tag{13}$$

We also define a symmetric $2n \times 2n$ matrix

$$P(\mu, \nu, \Gamma) = \begin{bmatrix} \nu K - K\Gamma - (K\Gamma)^T & X\Gamma - (X\Gamma)^T \\ (X\Gamma)^T - X\Gamma & \nu M + M\Gamma + (M\Gamma)^T + 2\mu\bar{C} \end{bmatrix} \tag{14}$$

for some scalars μ, ν, and some $n \times n$ matrix Γ. Finally, we define the $n \times 1$ matrix

$$N \equiv \frac{\partial}{\partial q} L(t_o, q^e, 0) + D(t_o, q^e, 0). \tag{15}$$

Notice that M, K, and $\bar{C}$ are symmetric, while X is skew-symmetric. If $2B - C$ is symmetric, then $X = 0$ and $P(\mu, \nu, \Gamma)$ is "matrix diagonal." By Corollary 2.1, $N = 0$ if the kinematic constraints are locally ineffective ($\mathcal{C}(t) \supset \mathcal{N}_h$ for all $t \geq t_o$ and some $h > 0$), which was assumed in Corollaries 3.1–3.2 of (Walker, 1988) and Corollary 3.3 of (Walker, 1992), but is not assumed here.

Corollary 4.1: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$:

(a) $\mathcal{N}_h \cap \mathcal{C}(t)$ is time invariant and the statement $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$ implies that $(q, 0) \in \mathcal{C}(t_o)$..

(b) L and D are time invariant, L is C^2-smooth, and $D(t_o, q, 0) = D(t_o, q^e, 0)$ for all $(q, 0) \in \mathcal{N}_h$.

(c) $u^T M u + (q - q^e)^T K (q - q^e) > 2N^T(q - q^e)$ for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$.

Then $(q^e, 0)$ is *stable.* If Q is pseudoconservative, then $(q^e, 0)$ is *not* asymptotically stable.

If in addition to (a)–(c),

(d) $(q^e, 0)$ is isolated, D is C^1-smooth, and $u^T C u > 0$ for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$, then $(q^e, 0)$ is *asymptotically* stable.

If the foregoing conditions (a), (b), and (d) hold, but in place of (c) we have its strong contradiction,

(e) For each $\epsilon > 0$ there exists some $(q, u) \in \mathcal{N}_\epsilon \cap \mathcal{C}(t_o)$ such that $u^T M u + (q - q^e)^T K (q - q^e) < 2N^T(q - q^e)$, then $(q^e, 0)$ is *unstable.*

Proof: Our current conditions (a)–(b) imply satisfaction of (a)–(b) in Theorem 3.1 and lead to the local estimates

$$|2G(q, u) - 2G(q^e, 0) - u^T M u - (q - q^e)^T K (q - q^e) + 2N^T(q - q^e)| \leq o(|q - q^e|^2 + |u|^2), \tag{16}$$

$$|D(t_o, q, u) - D(t_o, q^e, 0) + Cu| \leq o(|u|), \tag{17}$$

$$|u^T D(t_o, q, u) - u^T D(t_o, q^e, 0) + u^T \bar{C} u| \leq o(|u|^2), \tag{18}$$

for some continuous function $o(\cdot)$: $\mathcal{R} \to \mathcal{R}$ such that $o(x)/x \to 0$ as $x \to 0$. Hence, our current conditions (a)–(c) imply satisfaction of (c) in Theorem 3.1, our current (a)–(b) and (d) imply satisfaction of (d) in Theorem 3.1, and our current (a)–(b) and (e) imply satisfaction of (e) in Theorem 3.1. As all current conclusions follow from Theorem 3.1, the proof is complete. □

Corollary 4.2: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$ and some locally C^1-smooth function γ: $\mathcal{D} \to \mathcal{R}^n$:

(a) $\mathcal{N}_h \cap \mathcal{C}(t)$ is time invariant, $\gamma(q^e) = 0$, and the statement $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$ implies that both $(q, 0) \in \mathcal{C}(t_o)$, $(q, u + \gamma(q)) \in \mathcal{C}(t_o)$.

(b) L and D are time invariant, L is C^2-smooth, D is C^1-smooth, and $D(t_o, q, 0) = D(t_o, q^e, 0)$ for all $(q, 0) \in \mathcal{N}_h$.

(c) For every $\epsilon > 0$ there exists some $(q, u) \in \mathcal{N}_\epsilon \cap \mathcal{C}(t_o)$ for which $u^T M u + (q - q^e)^T K (q - q^e) < 2N^T(q - q^e)$.

(d) There exist real numbers $\mu \geq 0$, $\nu \geq 0$, such that

$$\begin{bmatrix} q - q^e \\ u \end{bmatrix}^T P(\mu, \nu, \Gamma) \begin{bmatrix} q - q^e \\ u \end{bmatrix} > 2N^T[\nu(q - q^e) - \gamma(q)]$$

for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$, where $\Gamma \equiv [(\partial/\partial q_j)\gamma_i(q^e)]$ (row i, column j). Then $(q^e, 0)$ is *unstable.*

Proof: Our current conditions (a)–(b) imply satisfaction of (a)–(b) in Theorem 3.2 and lead to the estimates (16)–(18). Hence, our current conditions (a)–(c) imply that (a)–(c) in Theorem 3.2 are satisfied. Since $\gamma(q^e) = 0$, we also obtain the estimates

$$|\gamma(q) - \Gamma(q - q^e)| \leq o(|q - q^e|), \tag{19}$$

$$|R(q, u) - u^T M \Gamma u - (q - q^e)^T (B\Gamma + \Gamma^T B - \Gamma^T C) u + (q - q^e)^T K \Gamma (q - q^e) - N^T \gamma(q)| \leq o(|q - q^e|^2 + |u|^2). \tag{20}$$

If we replace the nonunique pair (U, D) of Definition 2.1 with

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\hat{U}(t, q, u) \equiv U(t, q, u) + q^T(B - X)u,$$

$$\hat{D}(t, q, u) \equiv D(t, q, u) + (B - B^T - 2X)u,$$

we find that the corresponding

$$\hat{L}(t, q, u) = L(t, q, u) - q^T(B - X)u, \quad \hat{B} = X, \quad \hat{C} = \bar{C},$$

$$\hat{R}(q, u; \gamma) = R(q, u; \gamma) - u^T \frac{\partial}{\partial q}[(q - q^e)^T(B - X)\gamma(q)],$$

while G, Ω_h, N, M, and K are unchanged. For $\hat{R}$ the estimate (20) becomes

$$\left|\hat{R}(q, u; \gamma) - \tfrac{1}{2}\begin{bmatrix} q - q^e \\ u \end{bmatrix}^T P(0, 0, \Gamma)\begin{bmatrix} q - q^e \\ u \end{bmatrix} - (q - q^e)^T\Gamma^T\bar{C}u - N^T\gamma(q)\right| \leq o(|q - q^e|^2 + |u|^2). \quad (21)$$

Combining the estimates (16), (18), (21), we see that

$$\hat{R}(q, u; \gamma) - \mu u^T[\hat{D}(t_o, q, u) - \hat{D}(t_o, q^e, 0)] > \nu[G(q^e, 0) - G(q, u)]$$

for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$, some $h > 0$, provided that

$$\begin{bmatrix} q - q^e \\ u \end{bmatrix}^T P(\mu, \nu, \Gamma)\begin{bmatrix} q - q^e \\ u \end{bmatrix}^T - 2(q - q^e)^T\Gamma^T\bar{C}u > 2N^T[\nu(q - q^e) - \gamma(q)]$$

for all $(q, u) \in \mathcal{N}_h \cap \bar{\mathcal{C}}(t_o)$. If our current condition (d) is met for some $\mu = \mu_o$, this proviso is met for sufficiently large $\mu > \mu_o$. Since $\nu \geq 0$ and $G(q, u) < G(q^e, 0)$ for $(q, u) \in \Omega_h \cap \mathcal{C}(t_o)$, it follows that conditions (d)–(e) in Theorem 3.2 are met by $\hat{R}$ and G. Hence, $(q^e, 0)$ is unstable and the proof is complete. □

Although simpler than the general Theorems 3.1–3.2, Corollaries 4.1–4.2 may still be difficult to apply, depending upon the complexity of the kinematic constraints. In order to simplify further, we assume henceforth that all kinematic constraints are locally holonomic and schleronomic near $(q^e, 0)$; that is, for all $t \geq t_o$ and some $h > 0$,

$$\mathcal{N}_h \cap \mathcal{C}(t) = \left\{(q, u) \in \mathcal{N}_h \,|f(q) = 0 = \sum_{l=1}^{n} u_l \frac{\partial}{\partial q_l} f(q) \in \mathcal{N}^m\right\} \quad (22)$$

for some given function $f: \mathcal{O} \to \mathcal{R}^m$. We also assume that f is C^2-smooth and define several $m \times n$ matrices (row i, column j) as

$$F(q) \equiv \left[\frac{\partial}{\partial q_j} f_i(q)\right], \quad F_o \equiv F(q^e),$$

$$F_k \equiv \frac{\partial}{\partial q_k} F(q^e), \quad (23)$$

for $k = 1, 2, \ldots, n$. Notice that Eq. (22) becomes

$$\mathcal{N}_h \cap \mathcal{C}(t) = \{(q, u) \in \mathcal{N}_h \,|f(q) = 0 = F(q)u\}.$$

Recalling the $n \times 1$ matrix N of Eq. (15), we see that the equilibrium conditions of Corollary 2.1 are met if and only if $\delta^T N = 0$ for all $\delta \in \mathcal{R}^n$ such that $F_o\delta = 0$. Hence, there must exist at least one $w \in \mathcal{R}^m$ such that

$$F_o^T w = N. \quad (24)$$

Selecting any one such w, we define another $n \times n$ matrix Z whose ith row is $w^T F_i$,

$$Z \equiv \begin{bmatrix} w^T F_1 \\ \vdots \\ w^T F_n \end{bmatrix} = \sum_{k=1}^{m} w_k\left[\frac{\partial^2}{\partial q_i \partial q_j} f_k(q^e)\right] \text{ (row } i\text{, column } j\text{)}. \quad (25)$$

We notice that w and Z are unique if and only if F_o has rank m; moreover, $Z = Z^T$ and

$$w^T(\sum_{k=1}^{n}(q_k - q_k^e)F_k) = (q - q^e)^T Z \quad (26)$$

for all $q \in \mathcal{O}$.

Corollary 4.3: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$:

(a) $\mathcal{N}_h \cap \mathcal{C}(t)$ is described by (22) for all $t \geq t_o$.

(b) L and D are time invariant, L and f are C^2-smooth, and $D(t_o, q, 0) = D(t_o, q^e, 0)$ for all $(q, 0) \in \mathcal{N}_h$.

(c) $\delta^T M\delta > 0$ and $\delta^T(K + Z)\delta > 0$ for all $\delta \in \mathcal{R}^n$ such that $F_o\delta = 0$, $\delta \neq 0$. Then $(q^e, 0)$ is *stable*. If Q is pseudoconservative, then $(q^e, 0)$ is *not* asymptotically stable.

If in addition to (a)–(c),

(d) $(q^e, 0)$ is isolated, D is C^1-smooth, and $\delta^T C$-smooth, and $\delta^T C\delta > 0$ for all $\delta \neq 0 \in \mathcal{R}^n$ such that $F_o\delta = 0$, then $(q^e, 0)$ is *asymptotically* stable.

If the foregoing conditions (a), (b), and (d) hold, but in place of (c) we have its strong contradiction,

(e) $\delta^T(K + Z)\delta < 0$ for some $\delta \in \mathcal{R}^n$ such that $F_o\delta = 0$, then $(q^e, 0)$ is *unstable*.

Proof: Our current conditions (a)–(b) imply satisfaction of (a)–(b) in Corollary 4.1, and $F(q)u = f(q) = f(q^e) = 0 \in \mathcal{R}^m$ for all $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$. Since (23) implies

$$|F(q)u - F_o u| \leq o(|q - q^e| + |u|),$$

$$|f(q) - f(q^e) - F_o(q - q^e) - \tfrac{1}{2}\sum_{k=1}^{n}(q_k - q_k^e)F_k(q - q^e)| \leq o(|q - q^e|^2)$$

for all $(q, u) \in \mathcal{N}_h$, we find that

$$u^T F_o^T F_o u \leq o(|q - q^e|^2 + |u|^2), \quad (27)$$

$$(q - q^e)^T F_o^T F_o(q - q^e) \leq o(|q - q^e|^2), \quad (28)$$

$$|2N^T(q - q^e) + (q - q^e)^T Z(q - q^e)| = |2w^T F_o(q - q^e) + w^T\sum_{k=1}^{n}(q_k - q_k^e)^T F_k(q - q^e)| \leq o(|q - q^e|^2) \quad (29)$$

for all $(q, u) \in \mathcal{N}_h \cap \mathcal{C}(t_o)$, having employed Eqs. (24)–(26) in (29).

The estimates (27)–(29) show that condition (c) of Corollary 4.1 is met if both $M + \rho F_o^T F_o$ and $K + Z + \rho F_o^T F_o$ are positive definite for some $\rho \geq 0$, and this is equivalent to our current condition (c). These estimates also show that condition (e) of Corollary 4.1 is met if $K + Z + \rho F_o^T F_o$ is not positive semidefinite for any $\rho \geq 0$, and this is equivalent to our current (e).

The estimate (27) shows that condition (d) of Corollary 4.1 is met if $(q^e, 0)$ is isolated and $\bar{C} + \rho F_o^T F_o$ is positive definite

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

for some $\rho \geq 0$, and this is equivalent to our current condition (d). The proof is complete. □

Corollary 4.4: Suppose that $(q^e, 0)$ is a generalized equilibrium and the following conditions hold for some $h > 0$ and some locally C^1-smooth function γ: $\mathcal{O} \to \mathcal{R}^n$:

(a) $\mathcal{N}_h \cap \mathcal{C}(t)$ is described by (22) for all $t \geq t_o$; $\gamma(q^e) = 0$, and $F(q)\gamma(q) = 0$ for all $(q, 0) \in \mathcal{N}_h$.

(b) L and D are time invariant, L and f are C^2-smooth, D is C^1-smooth, and $D(t_o, q, 0) = D(t_o, q^e, 0)$ for all $(q, 0) \in \mathcal{N}_h$.

(c) $\Delta^T(K + Z)\Delta \prec 0$ for some $\Delta \in \mathcal{R}^n$ such that $F_o\Delta = 0$.

(d) For $\Gamma \equiv [(\partial/\partial q_j)\gamma_i(q^e)]$ (row i, column j) and some real numbers $\nu \geq 0$, $\mu \geq 0$,

$$\begin{bmatrix} \Delta \\ \delta \end{bmatrix}^T P(\mu, \nu, \Gamma) \begin{bmatrix} \Delta \\ \delta \end{bmatrix} + \Delta^T Z(\nu I - 2\Gamma)\Delta > 0$$

for all $(\Delta, \delta) \neq (0, 0) \in \mathcal{R}^n \times \mathcal{R}^n$ such that $F_o\Delta = 0 = F_o\delta$. Then $(q^e, 0)$ is *unstable.*

Proof: Our current conditions (a)–(b) imply satisfaction of conditions (a)–(b) in Corollary 4.2 and lead to the estimates (27)–(29). An argument made in the proof of Corollary 4.3 shows that condition (c) of Corollary 4.2 is met if our current condition (c) is met.

Since γ is C^1-smooth and $\gamma(q^e) = 0 \in \mathcal{R}^n$, (23) implies that

$$|F(q)\gamma(q) - F_o\gamma(q) - \sum_{k=1}^{n} (q_k - q_k^e)F_k\gamma(q)| \leq o(|q - q^e|^2)$$

for all $(q, u) \in \mathcal{N}_h$. Hence, our current condition (a) and the estimate (19) lead to

$$|F_o\gamma(q) + \sum_{k=1}^{n} (q_k - q_k^e)F_k\Gamma(q - q^e)| \leq o(|q - q^e|^2) \quad (30)$$

for all $(q, u) \in \mathcal{N}_h$. Using (30) with (24)–(26), we find that

$$|N^T\gamma(q) + (q - q^e)^T Z\Gamma(q - q^e)| = |w^T F_o\gamma(q) + w \sum_{k=1}^{n} (q_k - q_k^e)F_k\Gamma(q - q^e)| \leq o(|q - q^e|^2) \quad (31)$$

for all $(q, u) \in \mathcal{N}_h$.

The estimates (27)–(29) and (31) show that condition (d) of Corollary 4.2 is met if the $2n \times 2n$ matrix

$$P(\mu, \nu, \Gamma) + \begin{bmatrix} \rho F_o^T F_o + \nu Z - Z\Gamma - (Z\Gamma)^T & 0 \\ 0 & \rho F_o^T F_o \end{bmatrix}$$

is positive definite for some non-negative ρ, ν, μ, and this is equivalent to our current condition (d). The proof is complete. □

In Corollary 4.4 it seems best to choose $\nu = 0$ if $\Delta^T(K + Z)\Delta \leq 0$ for all Δ such that $F_o\Delta = 0$. If $\delta^T C\delta > 0$ for all $\delta \neq 0$ with $F_o\delta = 0$, then Corollary 4.4 should not be used since the instability result of Corollary 4.3 is simpler. In Corollary 4.4 the choice of a useful $\gamma(q)$ is not as difficult as it may appear to be. If F_o has rank m, we can determine $n-m$ linearly independent functions p_k: $\mathcal{O} \to \mathcal{R}^n$ such that $F(q)p_k(q) = 0$ for all $(q, 0) \in \mathcal{N}_h$, $k = 1, 2, n - m$. For any $(n - m) \times n$ real matrix $\hat{\Gamma}$, the choice

$$\gamma(q) \equiv [p_1(q) \ldots p_{n-m}(q)]\hat{\Gamma}(q - q^e)$$

meets condition (a) of Corollary 4.2, and the $n \times n$ matrix Γ in condition (d) is

$$\Gamma = [p_1(q^e) \ldots p_{n-m}(q^e)]\hat{\Gamma},$$

involving the $(n - m)n$ arbitrary parameters in $\hat{\Gamma}$. These parameters must then be chosen to satisfy condition (d) of Corollary 4.4. In particular, it is necessary that

$$\delta^T(K + Z)(\nu I - 2\Gamma)\delta > 0,$$

$$\delta^T[M(\nu I + 2\Gamma) + 2\mu C]\delta > 0,$$

for all $\delta \neq 0$ such that $F_o\delta = 0$. These requirements suffice for condition (d) if $X = 0$ in (13)–(14).

5 Example

Consider a slider/crank mechanism attached to a frame $\mathcal{F}$ motor-driven at a constant angular rate ω relative to the Earth $\mathcal{E}$, here assumed to be nearly inertial. See Figure 1. A small massless slider $\mathcal{B}_3$ moves along an Earth-fixed vertical axis, and is hinged at C to a uniform bar $\mathcal{B}_2$ (mass m_2, length l) hinged at its midpoint B to another uniform bar $\mathcal{B}_1$ (mass m_1, length $l\sqrt{2}$) hinged at A to the rotating frame $\mathcal{F}$. To make stability analysis difficult, we assume there is no friction at the slider or any of the hinges.

Choosing $q \equiv (\theta, \psi) \in \mathcal{R}^2 \equiv \mathcal{O}$ for the material collection $\mathcal{B}_1 + \mathcal{B}_2 + \mathcal{B}_3$, and defining $\bar{m} = m_1 + 2m_2$ for convenience, we find that

$$\mathcal{C}(t) = \left\{ (q, u) \in \mathcal{R}^2 \middle| \begin{matrix} 3 = 2\sqrt{2}\cos\theta + \cos\psi \\ 0 = 2\sqrt{2}u_\theta \sin\theta + u_\psi \sin\psi \end{matrix} \right\},$$

$$T(t, q, u) = \frac{1}{3}(m_1 + 3m_2)l^2(u_\theta^2 + \omega^2 \cos^2\theta) + \frac{1}{24} m_2 l^2(u_\psi^2 + \omega^2 \cos^2\psi) + \frac{9}{8}(m_1 + m_2)l^2\omega^2 - \frac{3\sqrt{2}}{4}\bar{m}l^2\omega^2 \cos\theta,$$

while Q is pseudoconservative ($D = 0$) with

$$\frac{\partial}{\partial q} L(t, q, u) = \frac{l}{24} \begin{bmatrix} -8(m_1 + 3m_2)l\omega^2 \sin 2\theta + 6\sqrt{2}\bar{m}(3l\omega^2 \sin\theta + 2g\cos\theta) \\ -m_2 l\omega^2 \sin 2\psi \end{bmatrix}.$$

$$U(t, q, u) = -(\bar{m}gl/\sqrt{2}) \sin\theta.$$

Defining $L \equiv T - U$, we see that neither U nor T is unique, L and $\mathcal{C}$ are time invariant, $\mathcal{C}$ is globally holonomic, and

For any ω, Corollary 2.1 implies that there exist at least two distinct generalized equilibria,

$$(q^e, 0)_1: \theta^e = \pi/4, \quad \psi^e = 0,$$

$$(q^e, 0)_2: \theta^e = -\pi/4, \quad \psi^e = 0,$$

but if ω^2 is larger than $\omega_c^2 \equiv 6\bar{m}g/(5m_1 + 8m_2)l$, there exist two more:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$(q^e, 0)_{3,4}: \psi^e \neq 0, \cos\psi^e = 3 - 2\sqrt{2}\cos\theta^e,$$

$$\omega^2 l(9m_1 + 24m_2 - 4\sqrt{2}(m_1 + 4m_2)\cos\theta^e)\tan\theta^e = -6\bar{m}g,$$

where $\theta_4^e = \theta_3^e$, $\psi_4^e = -\psi_3^e < 0$. Simple estimates show that $0 > \theta_3^e > -\pi/4$, $-\tan\theta_3^e > (\omega_c/\omega)^2$, $\theta_3^e \to -\pi/4$ as $\omega^2 \to \omega_c^2$, and $\theta_3^e \to 0$ as $\omega^2 \to \infty$.

At each $(q^e, 0)$ we see that the matrices of Eqs. (11)–(13) and (15) are

$$M = \frac{l^2}{12}\begin{bmatrix} 8(m_1 + 3m_2) & 0 \\ 0 & m_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = C = X,$$

$$K = \frac{l}{12}\begin{bmatrix} 8(m_1 + 3m_2)l\omega^2\cos 2\theta^e - 3\bar{m}\sqrt{2}(l\omega^2\cos\theta^e - 2g\sin\theta^2) & 0 \\ 0 & m_2 l\omega^2\cos 2\psi^e \end{bmatrix},$$

$$N = \frac{l}{24}\begin{bmatrix} -8(m_1 + 3m_2)l\omega^2\sin 2\theta^e + 6\sqrt{2}\bar{m}(3l\omega^2\sin\theta^e + 2g\cos\theta^e) \\ -m_2 l\omega^2\sin 2\psi^e \end{bmatrix},$$

while $\mathcal{N}_h \cap \mathcal{C}(t)$ is of the holonomic form (22) with $f(q) \equiv 3 - 2\sqrt{2}\cos\theta - \cos\psi$; hence, the matrices of Eqs. (23) are

$$F(q) = [2\sqrt{2}\sin\theta, \sin\psi], \quad F_o = [2\sqrt{2}\sin\theta^e, \sin\psi^e],$$

$$F_1 = [2\sqrt{2}\cos\theta^e, 0], \quad F_2 = [0, \cos\psi^e],$$

and from Eqs. (24)–(25) we find

$$24w = -4\sqrt{2}(m_1 + 3m_2)l^2\omega^2\cos\theta^e + 3\bar{m}l(3l\omega^2 + 2g\cot\theta^e),$$

$$Z = w\begin{bmatrix} 2\sqrt{2}\cos\theta^e & 0 \\ 0 & \cos\psi^e \end{bmatrix}.$$

We note that $12w = -m_2 l^2\omega^2\cos\psi^e$ if $\psi^e \neq 0$, conditions (a)–(b) of Corollary 4.3 are met, and $F_o\delta = 0$ if and only if

$$\delta = d(\sin\psi^e, -2\sqrt{2}\sin\theta^e) \in \mathcal{R}^2$$

for some scalar d.

For arbitrary $(g_1, g_2) \in \mathcal{R}^2$ the function

$$\gamma(q) \equiv \begin{bmatrix} \sin\psi \\ -2\sqrt{2}\sin\theta \end{bmatrix}[g_1 \vdots g_2]\begin{bmatrix} \theta - \theta^e \\ \psi - \psi^e \end{bmatrix}$$

satisfies condition (a) of Corollary 4.4 for each generalized equilibrium $(q^e, 0)$. This $\gamma(q)$ leads to

$$\Gamma = \begin{bmatrix} \sin\psi^e \\ -2\sqrt{2}\sin\theta^e \end{bmatrix}[g_1 \vdots g_2]$$

in condition (d) of Corollary 4.4.

At $(q^e, 0)_1$ we see that $\delta = (0, -2d)$, $\delta^T M\delta = m_2 l^2 d^2/3 \geq 0$,

$$24w = (5m_1 + 8m_2)l^2(\omega_c^2 + \omega^2) - 2m_2 l^2\omega,$$

$$\delta^T(K + Z)\delta = (m_2 l^2\omega^2 + 12w)d^2/3 \geq 0,$$

so Corollary 4.3 states that $(q^e, 0)_1$ is *stable,* but not asymptotically stable.

At $(q^e, 0)_2$ we see that $\delta = (0, 2d)$, $\delta^T M\delta = m_2 l^2 d^2/3 \geq 0$,

$$24w = (5m_1 + 8m_2)l^2(\omega^2 - \omega_c^2) - 2m_2 l^2\omega^2,$$

$$\delta^T(K + Z)\delta = (5m_1 + 8m_2)l^2(\omega^2 - \omega_c^2)d^2/3.$$

Hence, Corollary 4.3 states that $(q^e, 0)_2$ is *stable* (not asymptotically) for $\omega^2 > \omega_c^2$. Choosing $(g_1, g_2) = (0, 1)$ to specify $\gamma(q)$, we obtain

$$\Gamma = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix},$$

$$\delta^T M\Gamma\delta = 2m_2 l^2 d^2/3 \geq 0,$$

$$\delta^T(K + Z)\Gamma\delta = 2(5m_1 + 8m_2)l^2(\omega^2 - \omega_c^2)d^2/3.$$

Noting that $X = 0$ in (14) and choosing $\nu = 0$ in Corollary 4.4, we find that $(q^e, 0)_2$ is *unstable* for $\omega^2 < \omega_c^2$.

At both $(q^e, 0)_3$ and $(q^e, 0)_4$ we have $\omega^2 > \omega_c^2$, $\sin\theta^e < 0$, $\sin\psi^e \neq 0$. We find that

$$\delta^T M\delta = 2l^2((m_1 + 3m_2)\sin^2\psi^e + m_2\sin^2\theta^e)d^2/3 \geq 0,$$

$$\delta^T(K + Z)\delta = (k_{11} + w2\sqrt{2}\cos\theta^e)d^2\sin^2\psi^e + 8(k_{22} + w\cos\psi^e)d^2\sin^2\theta^e,$$

$$6(k_{11} + w2\sqrt{2}\cos\theta^e) = -4(m_1 + 3m_2)l^2\omega^2\sin^2\theta^e + 3\sqrt{2}\bar{m}gl/\sin\theta^e < 0,$$

$$12(k_{22} + w\cos\psi^e) = -m_2 l^2\omega^2\sin^2\psi^e < 0,$$

so condition (c) of Corollary 4.4 is met. Choosing $(g_1, g_2) = (0, 1)$ to specify $\gamma(q)$, we obtain

$$\Gamma = \begin{bmatrix} 0 & \sin\psi^e \\ 0 & -2\sqrt{2}\sin\theta^e \end{bmatrix},$$

$$\delta^T M\Gamma\delta = 4\sqrt{2}(-\sin\theta^e)((m_1 + 3m_2)\sin^2\psi^e + m_2\sin^2\theta^e)d^2/3 \geq 0,$$

$$\delta^T(K + Z)\Gamma\delta = 2\sqrt{2}(-\sin\theta^e)((k_{11} + w2\sqrt{2}\cos\theta^e)\sin^2\psi^e + 8(k_{22} + w\cos\psi^e)\sin^2\theta^e)d^2 \leq 0.$$

Noting that $X = 0$ in (14) and choosing $\nu = 0$ in Corollary 4.4, we find that both $(q^e, 0)_3$ and $(q^e, 0)_4$ are *unstable.*

Our results are summarized by the table of Fig. 2, but more can be said if there is sufficient viscous friction at the slider or any of the joints A, B, C to make the generalized force Q strongly pseudodissipative (see Definition 2.1). Then Theorem 3.1 shows that each claim of "stable" in Fig. 2 can be strength-

	$\omega^2 < \omega_c^2$	$\omega^2 > \omega_c^2$
$(q^e,0)_1$	stable	stable
$(q^e,0)_2$	unstable	stable
$(q^e,0)_{3,4}$	impossible	unstable

Fig. 2 Stability table for example

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ened to "asymptotically stable," and the "unstable" conclusions are unchanged. If there is sufficient viscous friction to make $\delta^T C\delta > 0$ for all $\delta \neq 0 \in \mathcal{R}^2$ with $F_o\delta = 0$, then this strengthening can be done *solely* on the basis of Corollary 4.3 and the more difficult Corollary 4.4 is not needed!

6 Concluding Remarks

The general stability results of Theorems 3.1–3.2 are intended for generally constrained systems, as were similar theorems in Walker (1988). For unconstrained systems the new theorems are almost the same as the old, but for constrained systems the new Theorems 3.1–3.2 are much improved; their conditions (a)–(b) are weaker. Theorems 3.1–3.2 of Walker (1988) produce none of the instability conclusions we obtained for the example in Section 5.

The simpler Corollaries 4.1–4.2 also are intended for generally constrained systems. They generalize the "simplest possible" corollaries obtained in Walker (1988, 1992) for unconstrained systems. Notice that if $\mathcal{C}(t) \supset \mathcal{N}_h$ for some $h > 0$ and all $t \geq t_o$, then $N = 0$ and $(q^e, 0)$ is isolated if $\det K \neq 0$; moreover, conditions (c) (d) (e) of Corollary 4.1 become definiteness requirements on M, K, and $\bar{C}$ as in Corollary 3.1 of Walker (1988). Also, condition (d) of Corollary 4.2 becomes a definiteness requirement on $P(\mu, \nu, \Gamma)$, which is $\hat{P}(\mu, \nu, \Gamma^T)$ of Corollary 3.3 in Walker (1992).

Most mechanisms are holonomically constrained, and often it is highly inconvenient to remove all holonomic constraints from the mathematical description, even if possible. In Section 2 we disputed the claim that this is always possible, and the example of Section 5 supports our argument. This example demonstrates that the simplified stability criteria of Corollaries 4.3–4.4 are quite useful for holonomically constrained systems.

If $(q^e, 0)$ is locally unconstrained and conditions (a)–(b) of Corollary 4.1 hold, the locally linearized motion equations are given by

$$M\ddot{x}(t) + (\bar{C} - 2X)\dot{x}(t) + Kx(t) = 0 \in \mathcal{R}^n,$$

where $x(t) \approx q(t) - q^e$ (Walker, 1988, 1992). Otherwise the "correct" linearization is much less obvious, often subject to dispute, and not discussed in textbooks on linear vibration. If the constraints are locally holonomic and scleronomic (see (22)), and if conditions (a)–(b) of Corollary 4.3 hold, the stability results in Section 4 indicate that the appropriate linearization is given by

$$M\ddot{x}(t) + (\bar{C} - 2X)\dot{x}(t) + (K + Z)x(t) = F_o^T\lambda(t) \in \mathcal{R}^n,$$

$$F_o x(t) = 0 \in \mathcal{R}^m,$$

where $\lambda(t) \in \mathcal{R}^m$ is an undetermined Lagrange multiplier. Stability conclusions for this particular linearization do not disagree with those of Corollaries 4.3–4.4 for the original (nonlinear) system.

Even for a "correct" linearization of the motion equations about a generalized equilibrium $(q^e, 0)$, stability conclusions *based* on this linearization can be invalid for the original nonlinear system. There are "critical cases" (present in the example of Section 5) for which linearized equations may yield incorrect stability conclusions. All conclusions in Sections 3–4 are exact, not based upon linearization.

All of the stability results in Sections 3–4 improve upon those in Walker (1988, 1992). The most apparently useful improvements are Corollaries 4.3–4.4, which provide very simple stability criteria for holonomically constrained nonlinear systems.

References

Kane, T. R., and Levinson, D. A., 1985, *Dynamics, Theory and Applications,* McGraw-Hill, New York.

Lagrange, J. L., 1788, *Mecanique Analytique,* new edition published by Mme. Ve. Courcier, Paris, 1811.

Likens, P. W., 1973, *Elements of Engineering Mechanics,* McGraw-Hill, New York.

Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems,* Plenum Press, New York.

Walker, J. A., 1988, "Pseudodissipative Systems I: Stability of Generalized Equilibria," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 3, pp. 681–686.

Walker, J. A., 1989, "Pseudodissipative Systems II: Stability of Reduced Equilibria," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, No. 2, pp. 466–473.

Walker, J. A., 1990, "Pseudodissipative Systems III: Global Behavior," *Quarterly of Applied Mathematics,* Vol. 48, No. 4, pp. 665–678.

Walker, J. A., 1992, "Instability of Generalized Equilibria of Pseudodissipative Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 59, No. 3, pp. 676–677.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Rotation of Clamped Parts due to Reciprocating Motion

E. S. Edelstein

J. J. Blech

Faculty of Mechanical Engineering,
Technion—Israel Institute of Technology,
Haifa 32000, Israel

Cylindrical and spherical parts, clamped in mechanical systems, rotate when the system is subjected to a dynamic reciprocating environmental condition. This rotation is a combination of the clamped part high "receptance" magnification near its natural frequencies, with frictional stick-slip at the clamping contact areas, when the system is excited to a reciprocating motion not collinear with the preloading direction. Such rotation may result in wear and fatigue failures in those systems, due to the rotation and the slip motion involved. Three cases are analyzed: cylindrical pins in hinges, balls in ball bearings, and spherical vessels clamped between spherical cavities.

1 Introduction

In mechanical systems, subjected to dynamic environmental conditions, rotation of cylindrical and spherical components assembled in between two surfaces occurs, due to the reciprocating motions in the system. Wear and fatigue failures may be a direct consequence of this phenomenon. This kind of assembly is widely used, for example, in systems with load carrying pins and rolling bearings. In rolling bearings rotation of a rolling elements may occur even in a nonrotating bearing. In addition, spherical vessels assembled in systems by clamping between surfaces, are found to change their position as a result of the systems' vibration. This rotation happens under relatively low excitation compared to the static forces acting on the parts. Moreover, if due to the "receptance" magnification, separation occurs, then no friction coefficient in the system will be able to prevent this rotation.

The rotation is a combination of the "receptance" magnification of the part, in the resilient clamping system, near its natural frequencies, with frictional stick-slip at the contact areas. This phenomenon occurs while the system is excited to reciprocating motion not collinear with the direction of the clamping forces at assembly. The dynamic force component in the preloading direction results in relative displacement between the part and its clamping system. At the same time, the other component, in the tangential direction (which is perpendicular to the preloading direction) is counteracted by friction forces at the contact areas. If instantaneously, the traction capability at one of the contact points becomes smaller than the driving tangential force component, developed as a result of the acceleration of the system, a slip displacement occurs. The higher traction capability at the opposite clamping side cannot balance the system as the displacement on this side changes into a rolling motion. For harmonic excitation, those small slip displacements alternate between the clamping sides, and accumulate into a net rotational displacement. A more detailed explanation and analysis of this phenomenon was given in Edelstein and Blech (1996).

The main parameters governing this rotational dynamic response are the natural frequencies of the part in the system which is acting as an elastic support, the damping coefficients, and the friction coefficients (static and dynamic) at the contact area. The elastic compliance of parts in contact in the normal direction was given by Hertz. The tangential elastic compliance of bodies in contact was analyzed by Mindlin (1949) and Mindlin and Deresiewicz (1953). Johnson had performed experiments to test the compliance and the energy dissipation of bodies in contact under oblique forces (Johnson, 1954, 1961). While the compliance in the normal direction is deterministic in general, the friction coefficients are arbitrary, the static friction coefficient changes at each slip onset and the dynamic coefficient of friction varies during slippage. Moreover, the compliance of bodies in contact in the tangential direction depends on the friction coefficient. In general, mean values of the friction coefficients lead to a good prediction of the dynamic rotational response.

A theoretical analysis of the aforementioned phenomenon is included in this paper, slippage onset conditions and an estimated analysis of the elastic compliance in such systems that govern the rotational response. Finally, to examine the analysis, an experimental test was performed. In the experiment, a perfect spherical ball was clamped between two spherical cavities. The system was excited to harmonic motion at various frequencies. In order to simulate this experiment by the theoretical model given here, additional experiments were performed to determine the natural frequencies, and estimate the damping and friction coefficients in the system.

2 Slippage Onset Conditions

Under static conditions, a ball, clamped between two surfaces, is loaded by two localized contact forces (N_1, N_2), normal to the surfaces, opposing each other and equal in magnitude (Fig. 1). When the system is brought into periodic linear motion, noncollinear with the preloading forces, friction forces (T_1, T_2) are developed between the ball and the clamping parts, at the contact points. Those forces balance the system to give coincident motion of the ball with the clamping parts.

If instantaneously, at one of the contact points, the traction capability is smaller than the friction force, necessary to balance the system for coincident motion, then a slip displacement occurs at this point.

Thus, at slippage onset either

$$|T_1| = |N_1\mu_s| \quad \text{or} \quad |T_2| = |N_2\mu_s|. \tag{1}$$

In order to establish the conditions for slippage onset in the system, the following assumptions are made:

(a) At steady-state low linear excitation the ball motion response is only translational (rotational response does not take place).

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, June 29, 1994; final revision, Mar. 3, 1995. Associate Technical Editor: J. T. Jenkins.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

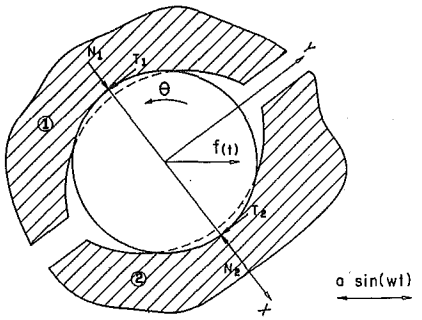


Fig. 1 The forces acting upon a ball clamped between two surfaces

(b) Although the elastic compliance for contact problems is not linear, in this case, a linear approximation can be adopted, because of the small displacements involved.

(c) Elastic displacement are accompanied by viscous damping.

For the elastic displacements (without slippage) the forces in the system are

$$N_1 = -K_n(\Delta_o + x) - C_n\dot{x}/2 \quad \text{and}$$

$$N_2 = -K_n(\Delta_o + x) - C_n\dot{x}/2$$

$$T_1 = -K_t y - C_t\dot{y}/2 \quad \text{and} \quad T_2 = -K_t y - C_t\dot{y}/2. \quad (2)$$

Where K_n, K_t are the elastic constants, and C_n, C_t are the damping coefficients in the x and y-directions. x, y are the instantaneous coordinates of the ball center in a coordinate system attached to the exciting system, and Δ_o is preloading elastic displacement at assembly.

For harmonic excitation with an acceleration amplitude a (having components a_x, a_y), the equations of motion are

$$m\ddot{x} + C_n\dot{x} + 2K_n x = -ma \sin(\omega t) \cos\alpha_1$$

$$m\ddot{y} + C_t\dot{y} + 2K_t y = -ma \sin(\omega t) \sin\alpha_1. \quad (3)$$

Thus, substituting Eqs. (2) into (3) we get

$$2N_1 + 2K_n\Delta_o = m\ddot{x} + ma \sin(\omega t) \cos\alpha_1 = ma_x \quad (4)$$

and

$$2T_1 = m\ddot{y} + ma \sin(\omega t) \sin\alpha_1 = ma_y, \quad (5)$$

where (a_x, a_y) in the right hand of those equations are the components of the acceleration of the ball in a Newtonian system.

Solving Eq. (3) leads to the following:
In the x direction,

$$a_x = A_n a_x \sin(\omega t - \theta_n)$$

$$A_n = \frac{a_x}{a_x} = \sqrt{\frac{1 + (2\zeta_n\omega/\Omega_n)^2}{[1 - (\omega/\Omega_n)^2]^2 + (2\zeta_n\omega/\Omega_n)^2}}$$

$$tg(\theta_n) = \frac{2\zeta_n(\omega/\Omega_n)^3}{1 - (\omega/\Omega_n)^2 + (2\zeta_n\omega/\Omega_n)^2}. \quad (6)$$

And, in the y direction,

$$a_y = A_t a_y \sin(\omega t - \theta_t)$$

$$A_t = \frac{a_y}{a_y} = \sqrt{\frac{1 + (2\zeta_t\omega/\Omega_t)^2}{[1 - (\omega/\Omega_t)^2]^2 + (2\zeta_t\omega/\Omega_t)^2}}$$

$$tg(\theta_t) = \frac{2\zeta_t(\omega/\Omega_t)^3}{1 - (\omega/\Omega_t)^2 + (2\zeta_t\omega/\Omega_t)^2}, \quad (7)$$

where Ω_n, Ω_t are the undamped natural frequencies, and ζ_n, ζ_t are the damping coefficients in the normal and tangential directions.

Due to symmetry, the slippage onset condition will be analyzed for contact area #1. After inserting Eq. (1) to (5) and dividing by Eq. (4) we get the static friction coefficient for slippage onset:

$$\mu_s = \text{Max} \left| \frac{A_t \sin(\omega t - \theta_t) \sin\alpha_1}{2K_n\Delta_o/(ma) - A_n \sin(\omega t - \theta_n) \cos\alpha_1} \right|, \quad (8)$$

where A_n and A_t are the amplifications of the excitation in the normal and tangential directions.

Also, defining the preloading force and the excitation inertial force in terms of multiples of the weight of the ball ($K_n\Delta_o = N_g\, mg$; $ma = N_a\, mg$). We get the condition for slippage onset:

$$\mu_s = \text{Max} \left| \frac{A_t \sin(\omega t - \theta_t) \sin\alpha_1}{2N_g/N_a - A_n \sin(\omega t - \theta_n) \cos\alpha_1} \right|. \quad (9)$$

The static friction coefficient necessary to prevent slippage on both sides at the same time is

$$\mu_s' = \frac{A_t \sin(\omega t - \theta_t) \sin\alpha_1}{2N_g/N_a + A_n \sin(\omega t - \theta_n) \cos\alpha_1}. \quad (10)$$

It can be seen from Eq. (9) that under certain conditions the denominator vanishes. Consequently, the static friction coefficient required for preventing rotation is infinite. This is the point of separation onset. If momentarily, due to the amplification of the excitation separation occurs, no friction coefficient will be able to prevent rotation.

In a special case, where the elastic compliance at the contact points is mainly in the normal direction (tangential compliance is negligible) the displacement of the ball, in the tangential direction, coincides with the clamping system as long as there is no slippage. Then Eq. (9) tends to

$$\mu_s = \text{Max} \left| \frac{\sin(\omega t) \sin\alpha_1}{2N_g/N_a - A_n \sin(\omega t - \theta_n) \cos\alpha_1} \right|. \quad (11)$$

In Fig. 2 the static friction coefficient necessary to prevent rotation as a function of the exciting frequency for three damping levels in the system. The damping levels were chosen so that in one case separation occurs ($\zeta = 0.015$) and in the other two, separation does not occur. Evidently, as seen from this figure, when there is separation between the clamped part and the clamping system, the static friction coefficient necessary to prevent rotation tends to infinity near the natural frequency.

However, if there is no separation, then the friction coefficient drops sharply near the natural frequency. This is a result of an approximately 90 deg phase difference between the excitation and the frequency response of the ball in the normal direction. The friction forces, in the tangential direction are synchronous with the excitation. In other words, the clamped part is at its maximum relative displacement in the normal direction when the system acceleration tends to zero. The zero acceleration results in no inertia forces in the tangential direction. Consequently, no friction forces are needed to give coincident motion of the system. Finally, when the normal force on one side reaches its minimum, due to maximum displacement, there are no traction forces required for coincident motion since there is no acceleration present at this instant. Near the natural frequency, although the normal force, on one side, does not reach its minimum, the necessity of friction forces for coincident motion results in the required friction coefficient to increase.

In general, when there is no separation, and the compliance at the contact areas is in both directions, normal and tangential, the friction coefficient is highly dependent on both natural frequencies. For example, Fig. 3 shows the friction coefficient μ_s necessary to prevent rotation (solid line), and the friction

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

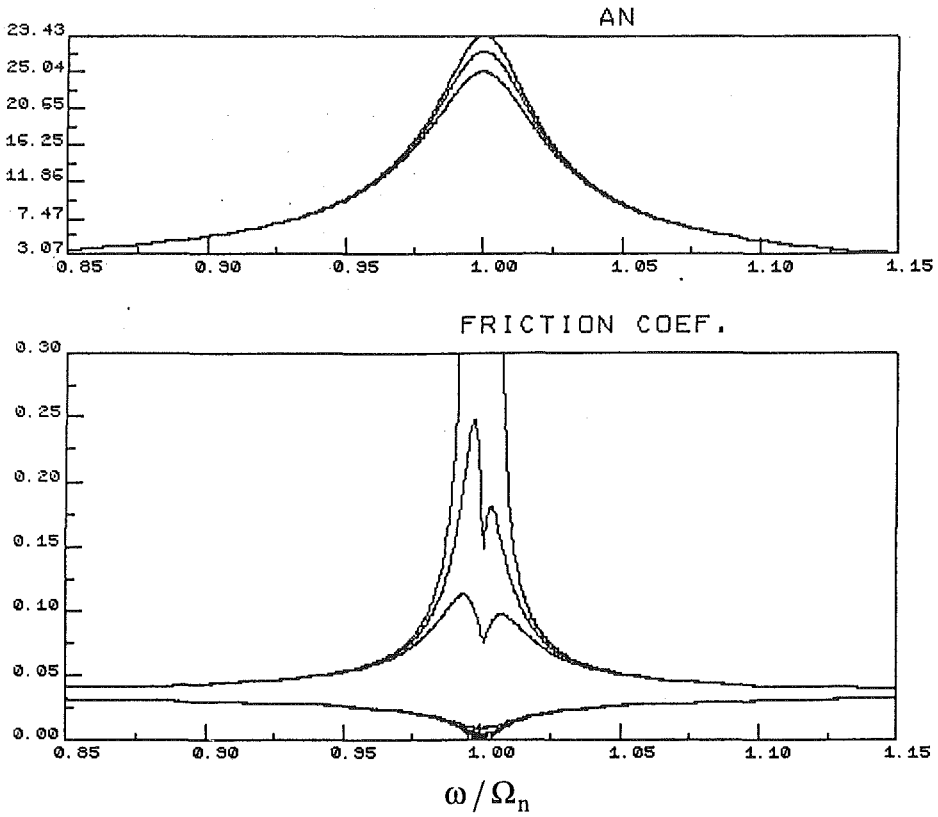


Fig. 2 Friction coefficient necessary to prevent rotation as a function of the exciting frequency for three damping levels (when the compliance at the contact point is mainly in the normal direction)

coefficient μ_s' necessary to prevent rotation on both sides (dashed line). However, when $\Omega_t \gg \Omega_n$ the friction coefficient necessary to prevent rotation increases near both natural frequencies. Near the natural frequency in the normal direction, the system will respond as in the case of a normal compliance only. However, near the natural frequency in the tangential direction the friction coefficient μ_s rises again, but this time μ_s' rises as well, eventually leading to slippage at both sides at the same time. This explains the periodic rotational response to excitation at a frequency near the natural frequency in the normal direction, and the nonperiodic rotational response to excitation near the natural frequency in the tangential direction, as introduced in Edelstein and Blech (1996).

As the tangential natural frequency decreases, and reaches the normal natural frequency, the slippage onset friction coefficient increases. Furthermore, the sharp decrease of the friction coefficient, at the natural frequency, does not repeat itself in this case. Therefore, when the natural frequencies in both directions are close to each other, the higher friction coefficient necessary to prevent rotation covers a wide range of excitation frequencies.

Finally, when $\Omega_t < \Omega_n$ the probability of rotational response to occur is mainly near the natural frequency in the tangential direction (at both sides at the same time).

Actually, as evident ahead, in most cases when contact compliance is involved, the natural frequencies in both directions are close to each other.

3 Analysis of the Elastic Compliance Governing the Rotation

In order to estimate the rotational response of a ball clamped between two surfaces, the equation of motion (Edelstein and Blech, 1995) has to be supplemented with the elastic compliance characteristic at the contact areas. When a ball is clamped between two cavities in two assumed to be semi-infinite bodies, the elastic parts comply, at the contact area, as a contact problem. In some cases, however, the compliance of the clamping parts can be a combination of the compliance as a contact problem and compliance of the clamping parts as a structure.

According to Hertz's solution, for rigid spherical part in contact with an elastic spherical cavity, the displacement α in the normal direction due to a normal load N is

$$\alpha = \left(\frac{9\pi^2}{16}\frac{N^2k^2}{1/R + 1/R_c}\right)^{1/3}; \quad k = \frac{1-\nu}{2\pi\mu} \tag{12}$$

where μ = modulus of rigidity, ν = Poison's ratio, R = balls radius, and R_c = clamping part radius.

Or, as for our problem, for a given displacement the normal force is

$$N = \frac{4}{3\pi}\frac{1}{k}\sqrt{\frac{(L-L_0)^3}{1/R + 1/R_c}} \tag{13}$$

where L_0 is the free centers' distance and L is the loaded centers' distance.

The elastic compliance in the tangential direction of two elastic bodies in contact was first studied by Mindlin (1949). He provides two solutions to the tangential compliance of two bodies in contact under a given normal force.

In the first case no slip at the contact area occurs. This solution leads to infinite stresses at the circumference of the contact area. Consequently, an infinite friction coefficient is necessary (the normal stresses, according to Hertz's theory, tends to zero at the circumference).

In his second solution, where traction limit is reached, slip occurs and the stresses in the tangential direction are the local traction limit (normal stresses multiplied by a friction constant). Presumably, this is a very sensitive assumption, thus, a stick-

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

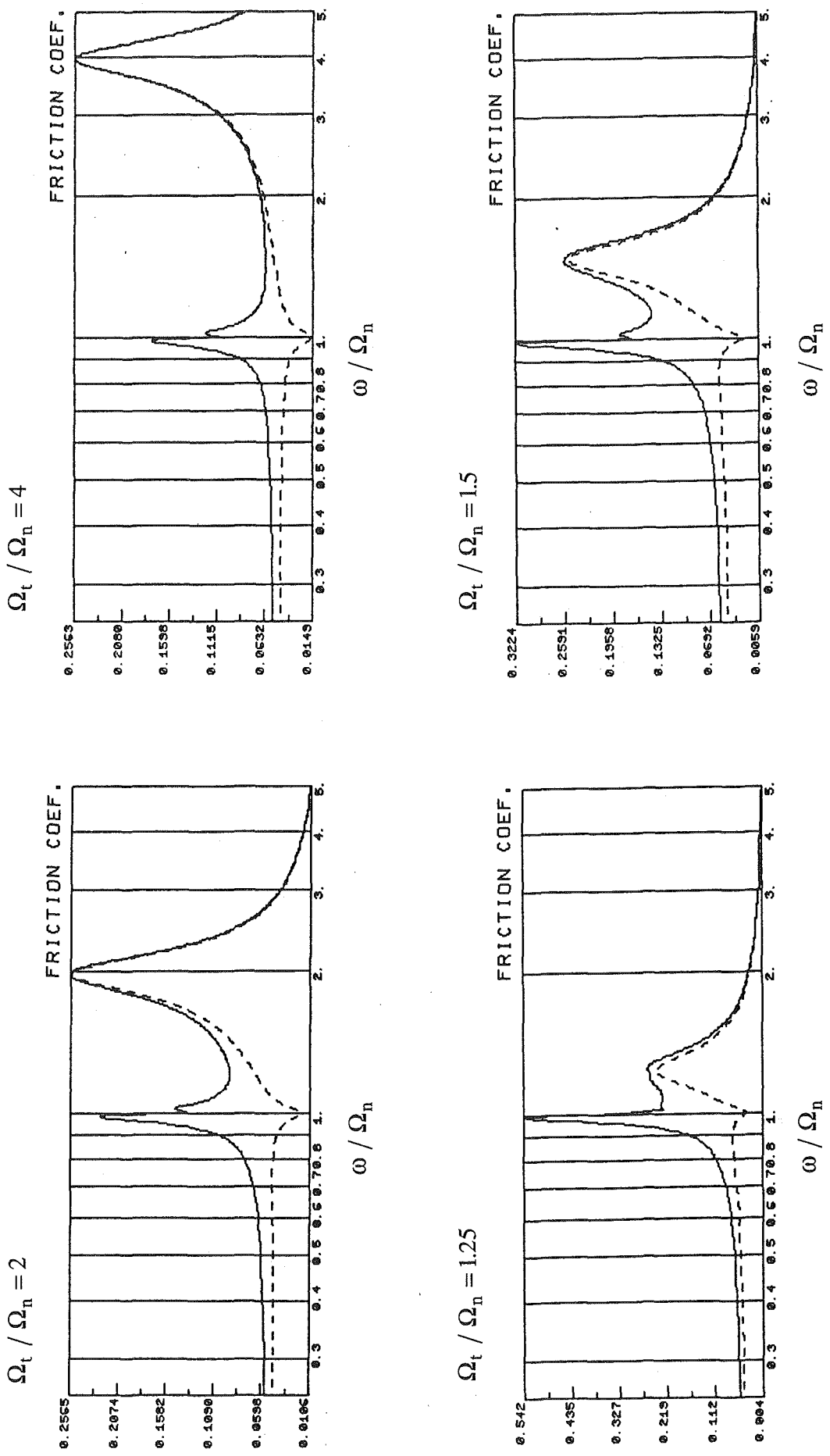


Fig. 3 The friction coefficient necessary to prevent rotation as a function of the exciting frequency for several Ω_t/Ω_n values ($\zeta_n = 0.01$; $\zeta_t = 0.07$)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

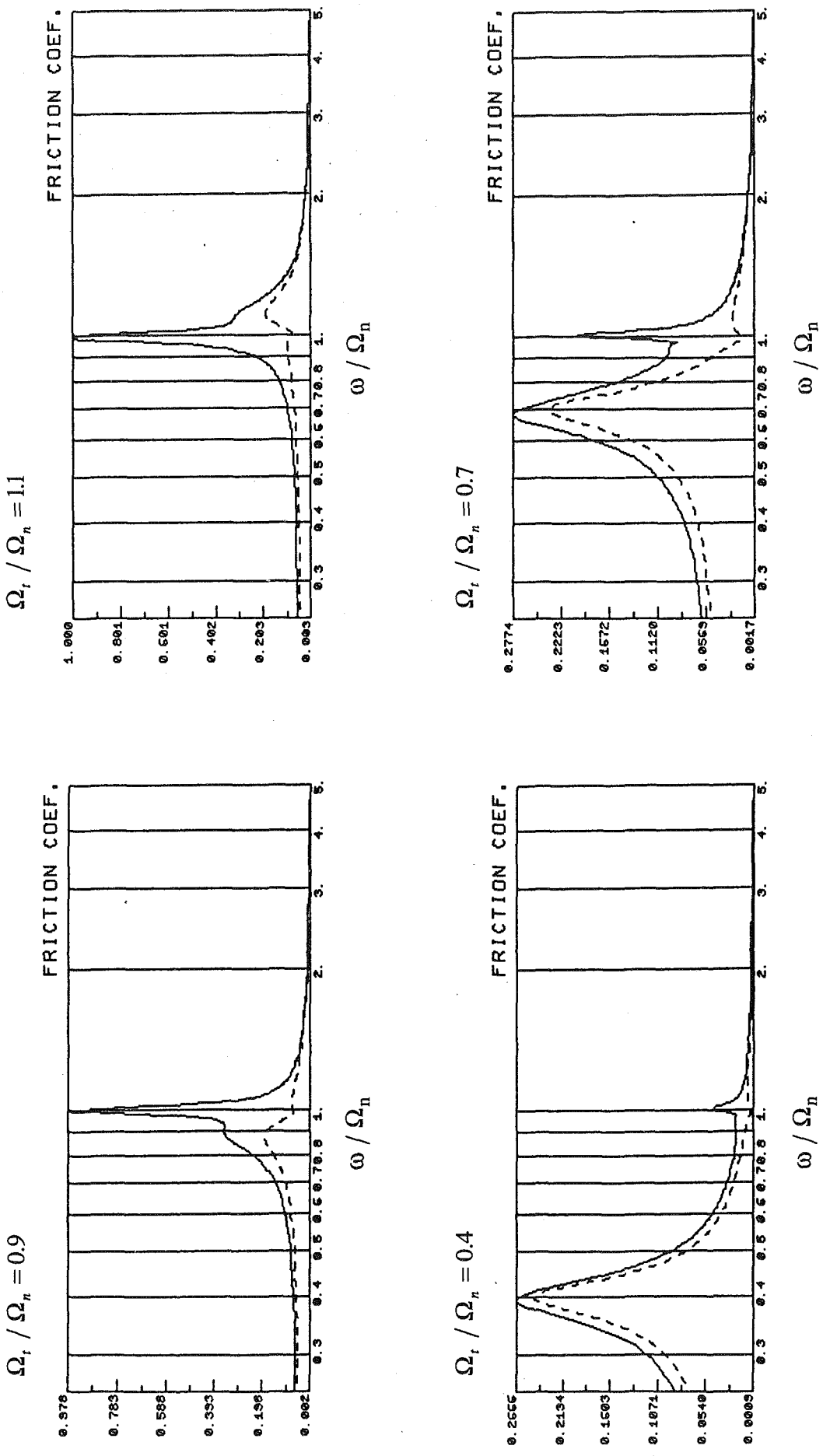


Fig. 3 (cont) The friction coefficient necessary to prevent rotation as a function of the exciting frequency for several Ω_t/Ω_n values ($\xi_n = 0.01$; $\xi_t = 0.07$)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

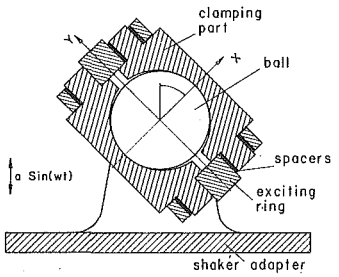


Fig. 4 Test apparatus—ball clamped between spherical surfaces

slip phenomenon can not occur at the contact area where traction limit is reached. However, under those assumptions Mindlin and Deresiewicz (1953) have built a theory for elastic spheres in contact under varying oblique forces.

In addition, Johnson (1954, 1961) performed experiments to test the surface interaction between elastically loaded bodies under tangential forces and energy dissipation at spherical surfaces in contact transmitting oscillating forces. In our case, for a wide range of force angles, the no slip condition assumption provides a good estimation for the dynamic response of the clamped part. Consequently, the elastic compliance in the tangential direction is

$$K_t = a\frac{8\mu}{2-\nu}. \tag{14}$$

In addition, according to Mindlin's analysis the ratio between elastic constants in the two directions is

$$\frac{K_t}{K_n} = \frac{2(1-\nu)}{2-\nu} = r. \tag{15}$$

For most engineering materials where ν is approximately 0.3, the elastic constant ratio is about 0.8, thus the natural frequencies ratio Ω_t/Ω_n is about 0.9. In general, when the relative elastic displacement between the ball and the clamping system is due to contact problem compliance, the natural frequencies in both directions are close to each other. Furthermore, for a ball clamped between "partially semi-infinite bodies," where the compliance of the clamping parts in the normal direction, is a combination of the compliance as a contact problem H_c and compliance of the clamping parts as a structure S_c, then the ratio between elastic constants in the two directions is

$$\frac{K_t}{K_n} = r\left(1+\frac{S_c}{H_c}\right). \tag{16}$$

Eventually this leads to an increase in the natural frequencies' ratio (Ω_t/Ω_n). To conclude, those natural frequencies can level for low structural compliance; and for large structural compliance, the natural frequencies' ratio can reach large values.

4 Ball Clamped Between Spherical Surfaces. Experiment and Simulation.

In order to test the theoretical analysis of the dynamic response of a ball clamped between spherical surfaces, and experiment was performed. The results of the experiment were simulated by numerical integration of the equations of motion (Edelstein and Blech, 1995) completed with the elastic constraints (Eqs. (13)–(16)).

A 75-mm radius aluminum ball was clamped between two "Perspex" parts with the cavities' radii of 77 mm (Fig. 4). The test apparatus was designed so that it was possible to change the angle of the preloading direction with respect to the excitation direction.

As a preparation to the main test, set-up experiments were performed:

(a) A static loading experiment to test the clamping part compliance, gave

$$N = 125500\delta[\text{mm}]^{1.34}\ [\text{N}].$$

(b) A dynamic experiment, in which the system was excited in the preloading direction and in the tangential direction to test the damping coefficients in the system, gave

$$\text{Viscous damping factors}\quad \zeta_n \approx 0.01;\quad \zeta_t \approx 0.07.$$

(c) A controlled torsion loading on the ball to estimate the friction coefficients (static and dynamic) at the contact areas, gave (mean; S.D.)

$$\mu_s = 0.225;\quad 0.07;\quad \mu_d = 0.153;\quad 0.012.$$

In one of the experiments (Fig. 5), the free distance between the clamping parts was smaller by 0.16 mm than the ball's diameter. Thus, after assembly, the ball was preloaded 90 times its weight. Results of this experiments and its simulation appear in Fig. 5.

The experimental test for the ball's rotational response was performed for a variety of preloading angles (with respect to the excitation). After verifying the governing natural frequencies in the test apparatus (Ω_n = 870 Hz; Ω_t = 760 Hz) the system was excited by a harmonic scanning 650-1100 Hz frequencies with ±9 g amplitude for three minutes. Those experiments were simulated with the characteristic values measured in the system (friction coefficients *mean values*). Clearly, in the test apparatus the friction coefficients (static and dynamic) are not at the mean value all the time and they vary every onset term and slippage (about 1500 times a second—about 50,000 times during experiment time). This is probably the reason for the difference between the total angular displacement in the experiment and its simulation. A research on the response of the system under random excitation and varying friction coefficients is now in progress. Although the test time was three minutes the simulation time is 30 seconds (ball's displacement linear to simulation time).

The main two conclusions from this experiments are (a) the ball rotates near the natural frequencies. For a higher excitation, the ball rotates in a wider range of frequencies. (b) The ball rotates in one direction and does not reveres its rotation (as in the case of a ball clamped between flat plates (Edelstein and Blech, 1995)). The unidirectional rotation is a result of the close natural frequencies. It is important to emphasize, that at 800 Hz the rotation is an accumulation of 1600 displacements per second, and each displacement is less than five micrometer (ball diameter 150 mm).

5 Conclusions

Axisymmetrical parts, symmetrically clamped between two surfaces, rotate due to linear reciprocating motion. This rotation can result from periodic or random excitation. For most design cases with clamped round and spherical parts (e.g., ball bearing, pins in hinges, or spherical vessels) the elastic compliance at the contact points, between the clamped part and the clamping parts, is a contact problem compliance. In those cases, the elastic compliance in both directions (normal and tangential) are close to each other, as derived from Mindlin's solution (Mindlin, 1949). As an example, for spherical parts with circular contact area the elastic constants' ratio is about 0.8. However, for cylindrical clamped parts with rectangular contact area this ratio is about 1, and for an elliptic contact area this ratio is between 0.7 ÷ 1 (depending on the direction of displacement relative to the ellipse axis). Consequently, the natural frequencies in both directions are closely related. Thus, in order to prevent this rotation to occur, due to relatively low excitation, high

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

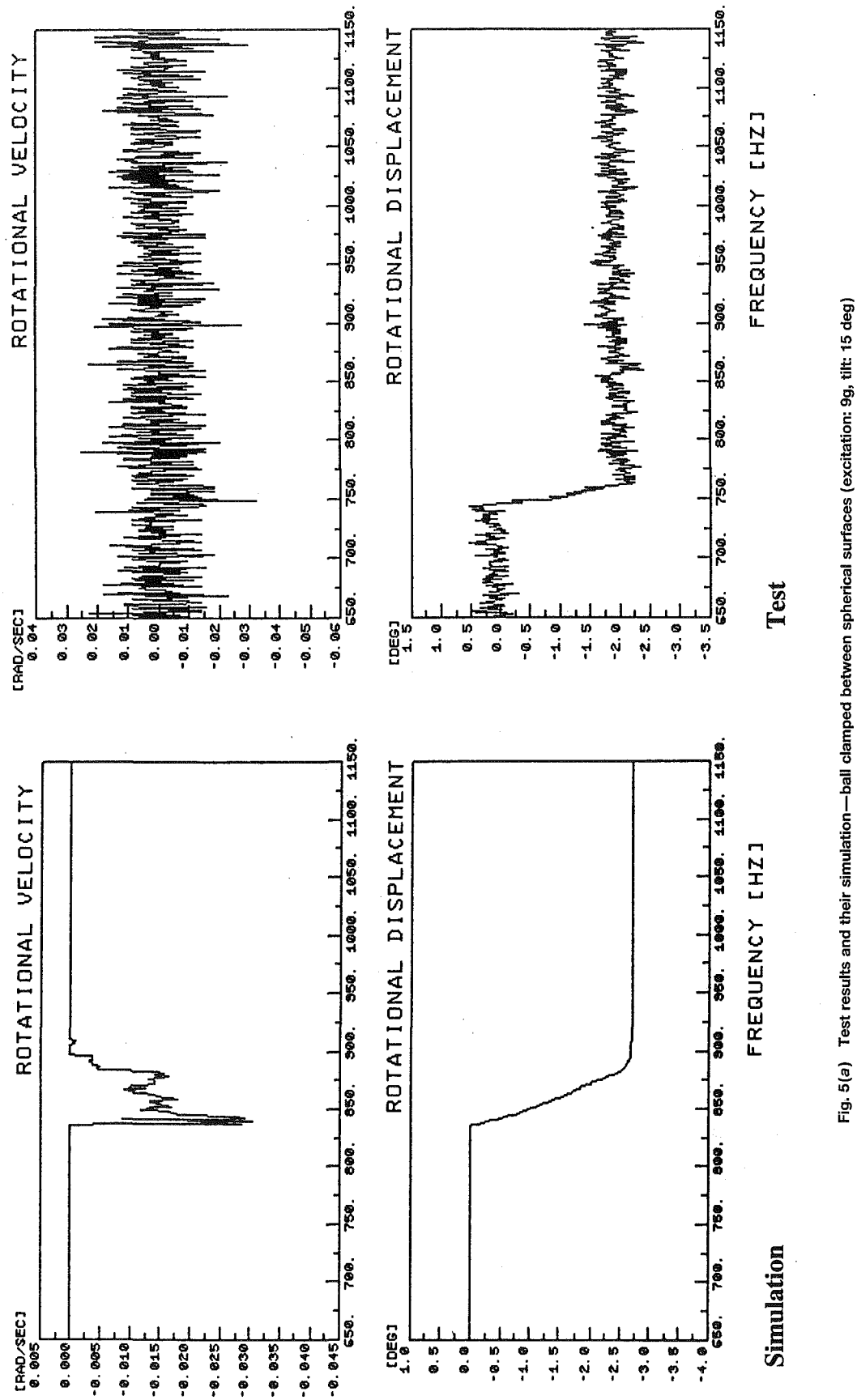


Fig. 5(*a*) Test results and their simulation—ball clamped between spherical surfaces (excitation: 9g, tilt: 15 deg)

friction coefficients are necessary. For example, in the experiment performed here, the ball was clamped by a force 90 times its own weight and rotated by excitation amplitude equivalent to nine times its weight. Furthermore, in this experiment the ball was clamped between plastic material parts where the damping coefficient is relatively high. However, for metal clamping parts

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

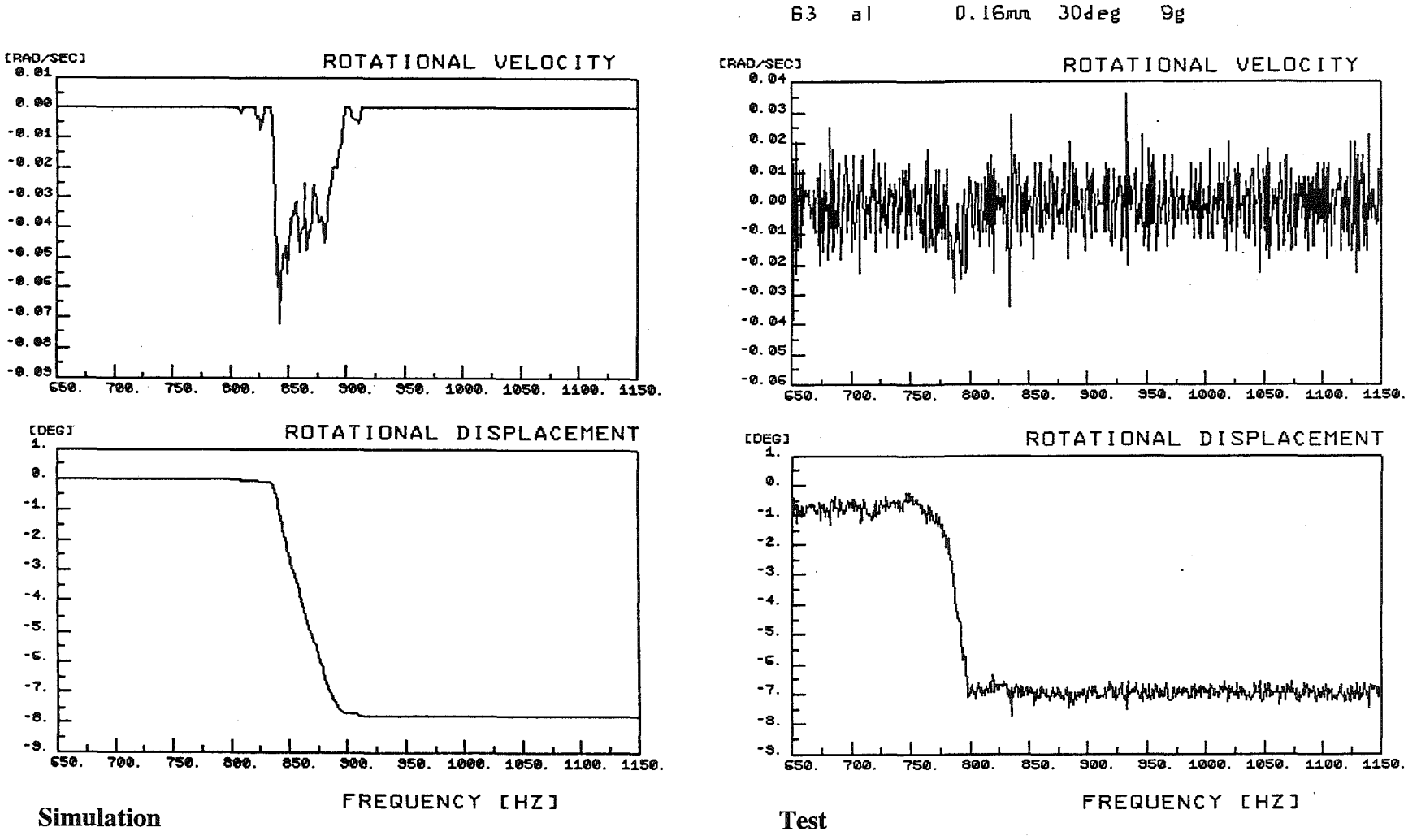


Fig. 5(*b*) Test results and their simulation—ball clamped between spherical surfaces (excitation: 9g, tilt: 30 deg)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

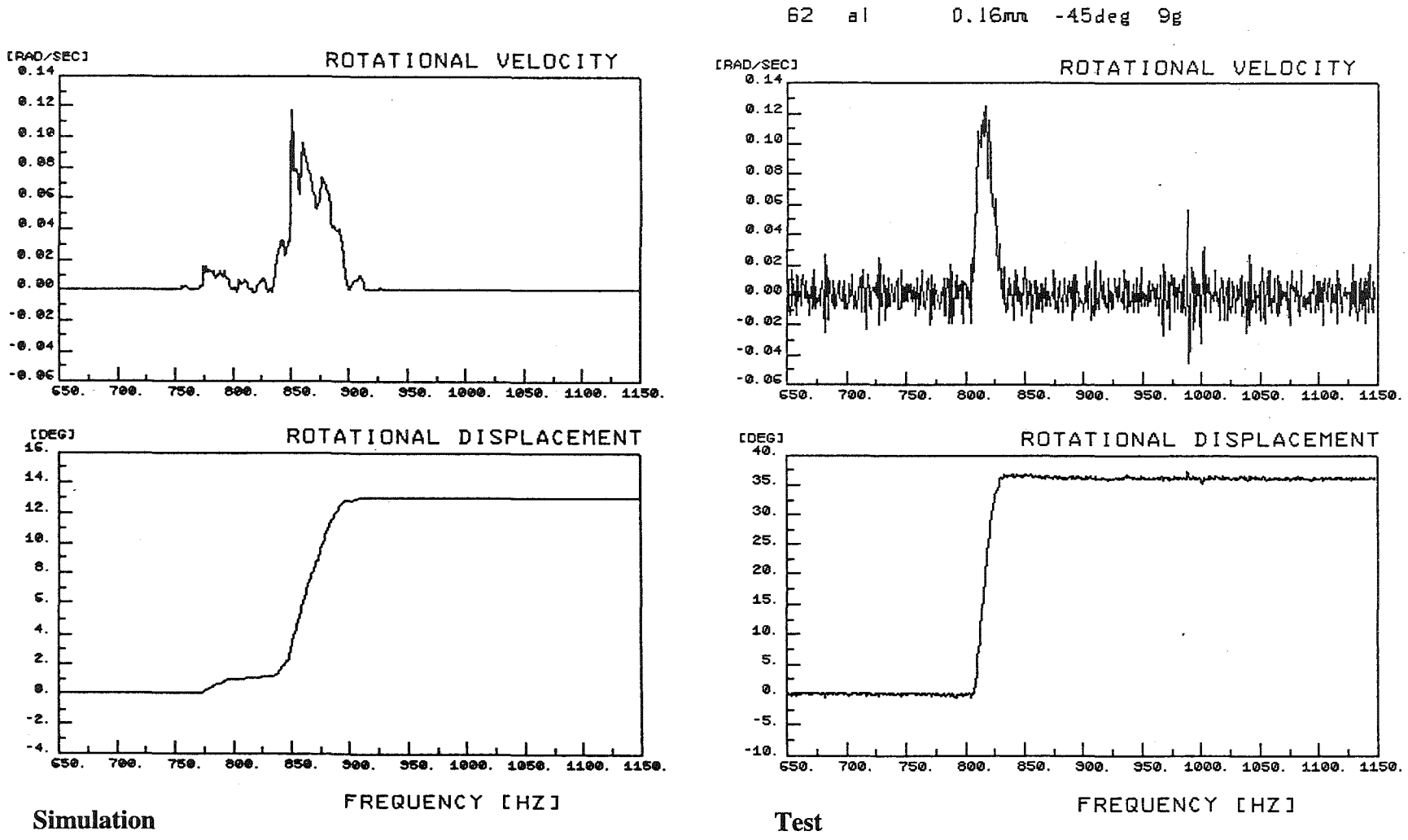


Fig. 5(c) Test results and their simulation—ball clamped between spherical surfaces (excitation: 9g, tilt: −45 deg)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

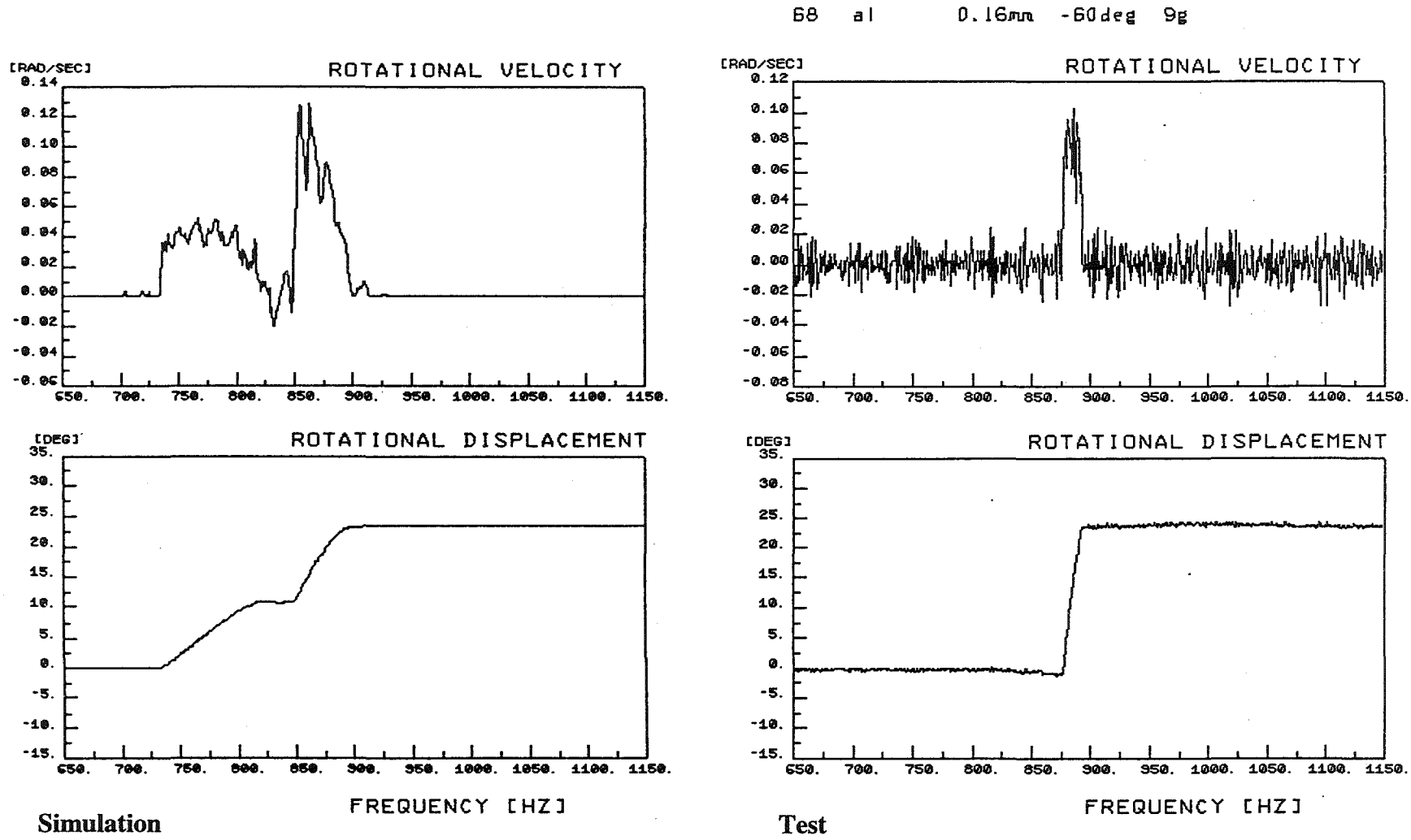


Fig. 5(*d*) Test results and their simulation—ball clamped between spherical surfaces (excitation: 9g, tilt: −60 deg)

where the damping coefficient is much lower this ratio, between the clamping forces and the excitation amplitude, can get even higher. This rotation will result in wear because of the slip motion involved. Such rotation may develop, under dynamic environmental conditions, even in nonrotating rolling bearings. Furthermore, cylindrical pins, which are widely used, for example as an attachment device between hanging equipment and airplane wings (naturally loaded between two surfaces), can be rotated by means of the environmental conditions of the aircraft.

References

Edelstein, E. S., and Blech, J. J., 1996, "Rotation of a Clamped Spherical Ball Due to Linear Reciprocating Motion," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp.

Johnson, K. L., 1954, "Surface Interaction Between Elastically Loaded Bodies Under Tangential Forces," *Proc. Roy. Soc.*, London, Vol. A230.

Johnson, K. L., 1961, "Energy Dissipation at Spherical Surfaces in Contact Transmitting Oscillating Forces," *Journal Mechanical Engineering Science,* Vol. 3, pp. 362–368.

Johnson, K. L., 1982, "One Hundred Years of Hertz Contact," *Proc. Instn. Mech. Engrs.,* Vol. 196, pp. 363–378.

Korycki, J., 1979, "Mathematical Model of the Stick-Slip Phenomenon," *Wear,* Vol. 55, pp. 261–263.

Mindlin, R. D., 1949, "Compliance of Elastic Bodies in Contact," ASME JOURNAL OF APPLIED MECHANICS, pp. 259–268.

Mindlin, R. D., and Deresiewicz, H., 1953, "Elastic Spheres in Contact Under Varying Oblique Forces," ASME JOURNAL OF APPLIED MECHANICS, pp. 327–344.

Rowson, D. M., 1975, "An Analysis of Stick-Slip Motion," *Wear,* Vol. 31, pp. 213–218.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

X. Markenscoff
Fellow ASME.

Luqun Ni

Department of Applied Mathematics
and Engineering Sciences,
University of California, San Diego,
La Jolla, CA 92093-0411

The Debonded Interface Anticrack[1]

The debonded interface anticrack is treated analytically and the singularity at the tip is found to vary between $\frac{1}{2}$ and 1, with the dependence on the material constants combination explicitly obtained. While the case of uniform tension and shear loading at infinity has been solved, the method of solution, which consists of distributing dislocation and line load densities, readily lends itself to solution for other point loadings, such as concentrated forces or dislocations.

Introduction

The solution is obtained for a rigid line inclusion (named "anticrack" by Dundurs and Markenscoff, 1989) which lies at the interface between two isotropic half-spaces, and which is debonded, so that a crack is formed on the one side of the rigid inclusion (Fig. 1). While a single crack, or a single line inclusion (Ballarini, 1990; Wang, et al., 1985; Wu, 1990), at an interface produce $r^{-1/2}$ singular fields near their tips, the combination of crack inclusion produces a singularity at the tip which varies from $-\frac{1}{2}$ to -1 according to the relevant softness of the two half-spaces. This dependence is obtained analytically, and interpreted physically. On the same subject Dempsey and Sinclair (1981) have considered an infinite wedge with different boundary conditions, including the clamped-free one. The eigenvalues that correspond to that problem should coincide with the debonded anticrack ones for the crack-tip singularity; however, no explicit values were presented by Dempsey and Sinclair. The anisotropic wedge of two different materials was subsequently treated by Ting (1986) and the case of semi-infinite crack, with free-fixed faces, was also considered for anisotropic and isotropic solids and the singularities obtained. The method used here is not only valid for the constant stress at infinity case as presented, but can be easily extended to other loadings, including point force and dislocation loadings as in Markenscoff et al. (1994).

1 Formulation of the Governing Equations

From the mathematical point of view, this problem is a mixed boundary value problem, with stress-free boundary conditions on one side of the cut at the bimaterial surface, and displacement boundary conditions at the other. Mixed boundary value problems like this may be treated by complex variables and singular integral equations techniques as was done for a homogenes medium by Sherman (1940), Muskhelishvili (1953), Keer (1975), and Erdogan (1969, 1965).

The method adopted here is based on distributing both dislocations and line loads on the cut surface, since crack problems may be solved by distributing dislocations and rigid line inclusion problems by distributing line loads. This method has the added advantage that it gives an insight into the physics of the problem. Indeed a crack, absorbs a dislocation, an r^{-1} singularity, when it approaches a crack-tip, while it does not absorb the concentrated force, also an r^{-1} singularity. For anticracks, considering them as a distribution of line loads, also corresponds to the physics of the problem: indeed, dually, the anticrack absorbs the concentrated force when it approaches its tip, while it does not absorb the dislocation, both r^{-1} singularities (Dundurs, Markenscoff, 1989).

For the debonded anticrack the method of solution is based on the distribution of density of line loads p_x, p_y and dislocations b_x, b_y in the interval $(-a, a)$. The stress and deformation gradients produced by each of these point "loads" constitute the Green's function for the problem at hand. Below we give the Green's functions for each of p_x, p_y, b_x, b_y (some of them already given by Markenscoff, Ni, and Dundurs (1994)). The coordinate system is shown in Fig. 1.

(1) Due to $p_x(\xi)$ at the interface:

$$\sigma_{xx}(0^-, y) = \frac{c}{4\mu_1}[\kappa_1(1-\beta) + (1+\beta)]p_x(\xi)\delta(y-\xi)$$

$$\sigma_{xy}(0^-, y) = \frac{-c}{4\pi\mu_1}[\kappa_1(1-\beta) - (1+\beta)]p_x(\xi)\frac{1}{y-\xi}$$

$$\sigma_{xx}(0^+, y) = \frac{-c}{4\mu_2}[\kappa_2(1+\beta) + (1-\beta)]p_x(\xi)\delta(y-\xi)$$

$$\sigma_{xy}(0^+, y) = \frac{-c}{4\pi\mu_2}[\kappa_2(1+\beta) - (1-\beta)]p_x(\xi)\frac{1}{y-\xi}$$

$$\frac{\partial u_y}{\partial y}(0, y) = \frac{c}{8\mu_1\mu_2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)]p_x(\xi)$$

$$\frac{\partial u_x}{\partial y}(0, y) = \frac{-c}{8\mu_1\mu_2\pi}[\kappa_2(1+\beta) + \kappa_1(1-\beta)]\frac{p_x(\xi)}{y-\xi}. \quad (1)$$

(2) Due to $p_y(\xi)$ at the interface:

$$\sigma_{xx}(0^-, y) = \frac{c}{4\pi\mu_1}[\kappa_1(1-\beta) - (1+\beta)]p_y(\xi)\frac{1}{y-\xi}$$

$$\sigma_{xy}(0^-, y) = \frac{c}{4\mu_1}[\kappa_1(1-\beta) + (1+\beta)]p_y(\xi)\delta(y-\xi)$$

$$\sigma_{xx}(0^+, y) = \frac{c}{4\pi\mu_2}[\kappa_2(1+\beta) - (1-\beta)]p_y(\xi)\frac{1}{y-\xi}$$

$$\sigma_{xy}(0^+, y) = \frac{-c}{4\mu_2}[\kappa_2(1+\beta) + (1-\beta)]p_y(\xi)\delta(y-\xi)$$

$$\frac{\partial u_y}{\partial y}(0, y) = \frac{-c}{8\pi\mu_1\mu_2}[\kappa_2(1+\beta) + \kappa_1(1-\beta)]p_y(\xi)\frac{1}{y-\xi}$$

[1] This paper is dedicated to Professor John Dundurs.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Jan. 17, 1995; final revision, June 28, 1995. Associate Technical Editor: L. T. Wheeler.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

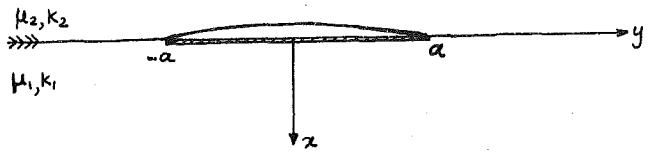


Fig. 1 The interface anticrack debonded as the upper side

$$\frac{\partial u_x}{\partial y}(0, y) = \frac{-c}{8\mu_1\mu_2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)]p_y(\xi)\delta(y-\xi). \quad (2)$$

(3) Due to $b_x(\xi)$ at the interface:

$$\sigma_{xy}(0^-, y) = \sigma_{xy}(0^+, y) = c\beta b_x(\xi)\delta(y-\xi)$$

$$\sigma_{xx}(0^-, y) = \sigma_{xx}(0^+, y) = \frac{-c}{\pi} b_x(\xi)\frac{1}{y-\xi}$$

$$\frac{\partial u_y}{\partial y}(0^-, y) = \frac{c}{4\pi\mu_2}[(1-\beta) - \kappa_2(1+\beta)]b_x(\xi)\frac{1}{y-\xi}$$

$$\frac{\partial u_x}{\partial y}(0^-, y) = \frac{-c}{4\mu_1}[\kappa_2(1+\beta) + (1-\beta)]b_x(\xi)\delta(y-\xi)$$

$$\frac{\partial u_y}{\partial y}(0^+, y) = \frac{c}{4\pi\mu_1}[(1+\beta) - (1-\beta)\kappa_1]b_x(\xi)\frac{1}{y-\xi}$$

$$\frac{\partial u_x}{\partial y}(0^+, y) = \frac{c}{4\mu_1}[(1+\beta) + (1-\beta)\kappa_1]b_x(\xi)\delta(\xi-y). \quad (3)$$

(4) Due to $b_y(\xi)$ at the interface:

$$\sigma_{xy}(0^-, y) = \sigma_{xy}(0^+, y) = -\frac{c}{\pi} b_y(\xi)\frac{1}{y-\xi}$$

$$\sigma_{xx}(0^-, y) = \sigma_{xx}(0^+, y) = -c\beta b_y(\xi)\delta(y-\xi)$$

$$\frac{\partial u_y}{\partial y}(0^-, y) = -\frac{c}{4\mu_1}[\kappa_2(1+\beta) + (1-\beta)]b_y(\xi)\delta(y-\xi)$$

$$\frac{\partial u_x}{\partial y}(0^-, y) = -\frac{c}{4\pi\mu_2}[(1-\beta) - \kappa_2(1+\beta)]b_y(\xi)\frac{1}{y-\xi}$$

$$\frac{\partial u_y}{\partial y}(0^+, y) = \frac{c}{4\mu_1}[(1+\beta) + (1-\beta)\kappa_1]b_y(\xi)\delta(y-\xi)$$

$$\frac{\partial u_x}{\partial y}(0^+, y) = -\frac{c}{4\pi\mu_1}[(1+\beta) - (1-\beta)\kappa_1]b_y(\xi)\frac{1}{y-\xi}. \quad (4)$$

Among the elastic constants μ_1, κ_1, μ_2, κ_2 and the Dundurs constants α, β (Dundurs, 1969) the following relations hold true:

$$\alpha = \frac{\mu_2(\kappa_1+1) - \mu_1(\kappa_2+1)}{\mu_2(\kappa_1+1) + \mu_1(\kappa_2+1)}$$

$$\beta = \frac{\mu_2(\kappa_1-1) - \mu_1(\kappa_2-1)}{\mu_2(\kappa_1+1) + \mu_1(\kappa_2+1)}$$

$$m \equiv \frac{\mu_2(1+\kappa_1)(1-\beta^2)}{1-\alpha} = \frac{\mu_1(1+\kappa_2)(1-\beta^2)}{1+\alpha}$$

$$c \equiv \frac{2\mu_1\mu_2}{m}$$

$$\frac{c}{\mu_1} = \frac{2(1+\alpha)}{(1+\kappa_1)(1-\beta^2)}, \quad \frac{c}{\mu_2} = \frac{2(1-\alpha)}{(1+\kappa_2)(1-\beta^2)} \quad (5)$$

where the Kolosov constant κ is $\kappa = 3 - 4\nu$ for plane strain, $\kappa = (3-\nu)/(1+\nu)$ for plane stress, and ν denotes the Poisson's ratio.

Let us consider now the problem of two bonded half-spaces with a debonded rigid line inclusion at the interface (as in Fig. 1) loaded by uniform loading T_x, T_y, T_{xy} at infinity. While this solution will be presented here, solving the problem with a different loading, such as a concentrated force or dislocation, can be achieved easily by this method, as for instance in the case of a rigid line inclusion at an interface (Markenscoff, Ni, and Dundurs 1994).

The distribution of line loads and dislocation densities must be determined from the boundary condition on the faces of the debonded anticrack.

Stress-free condition at $x = 0^-$:

$$\sigma_{xx} = 0, \quad \sigma_{xy} = 0, \quad -a < y < a \quad (6)$$

Rigid body displacement condition at $x = 0^+$:

$$e_{yy} = 0, \quad \Delta k = \frac{\partial u_x}{\partial y} = 0, \quad -a < y < a \quad (7)$$

where e_{yy} denotes the yy-component of the strain, and Δk denotes the change in the curvature. Substituting into (6) and (7) the contributions due to the distribution of dislocations, line loads (as obtained from (1) ~ (4)) and external loading $[T_x, T_y, T_{xy}]$, the boundary conditions (6) yield the integral equations:

$$(6a): \quad \frac{-\mu_1}{\pi}\left[\pi\beta b_y + \frac{1}{y} * b_x\right] + \frac{1}{4}[\kappa_1(1-\beta) + (1+\beta)]p_x + \frac{1}{4\pi}[\kappa_1(1-\beta) - (1+\beta)]\frac{1}{y} * p_y = \frac{-T_x(1+\kappa_1)(1-\beta^2)}{2(1+\alpha)} \equiv -\hat{T}_x \quad (8a)$$

$$(6b): \quad \mu_1\beta b_x - \frac{\mu_1}{\pi}\frac{1}{y} * b_y - \frac{1}{4\pi}[\kappa_1(1-\beta) - (1+\beta)]\frac{1}{y} * p_x + \frac{1}{4}[\kappa_1(1-\beta) + (1+\beta)]p_y = -\frac{T_{xy}(1+\kappa_1)(1-\beta^2)}{2(1+\alpha)} \equiv -\hat{T}_{xy}. \quad (8b)$$

Also, after substitution of the contributions of dislocations and line loads according to (1) ~ (4), the boundary conditions (7) yield the integral equations:

$$(7a): \quad \frac{\mu_2}{\pi}[(1+\beta) - (1-\beta)\kappa_1]\frac{1}{y} * b_x + \mu_2[(1+\beta) + (1-\beta)\kappa_1]b_y + \frac{1}{2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)]p_x - \frac{1}{2\pi}[\kappa_2(1+\beta) + \kappa_1(1-\beta)]\frac{1}{y} * p_y = -\frac{2\epsilon_{yy}(1+\kappa_1)(1-\beta^2)\mu_2}{(1+\alpha)} \equiv -\hat{\epsilon}_{yy} \quad (8c)$$

$$(7b): \quad \mu_2[(1+\beta) + (1-\beta)\kappa_1]b_x - \frac{\mu_2}{\pi}[(1+\beta) - (1-\beta)\kappa_1]\frac{1}{y} * b_y - \frac{1}{2\pi}[\kappa_2(1+\beta)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+ \kappa_1(1-\beta)]\frac{1}{y} * p_x - \frac{1}{2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)]p_y$$

$$= -\frac{\tilde{\omega}}{2m} \equiv -\hat{\omega} \quad (8d)$$

where ϵ_{yy} is due to the external loading and $\tilde{\omega} \equiv (\partial u_x^e/\partial y) - \omega$, with $\partial u_x^e/\partial y$ due to external loading and ω denoting the rotation of the anticrack. The symbol $(1/y) * b$ denotes the convolution $\int_{-a}^{a} (b(\zeta))/(y-\zeta)d\zeta$.

The system of Eqs. (8) can be then written in the form

$$\frac{1}{\pi} H_1\left(\frac{1}{y} * \hat{p}\right) + H_2\hat{p} = \hat{T} \quad (9)$$

where

$$\hat{p} = \begin{pmatrix} b_x \\ b_y \\ p_x \\ p_y \end{pmatrix}, \quad \hat{T} = \begin{pmatrix} \hat{T}_x \\ \hat{T}_{xy} \\ \hat{\epsilon}_{yy} \\ \hat{\omega} \end{pmatrix},$$

$$H_1 = \begin{bmatrix} -\mu_1 & 0 & 0 & \frac{1}{4}[\kappa_1(1-\beta) - (1+\beta)] \\ 0 & \mu_1 & \frac{1}{4}[\kappa_1(1-\beta) - (1+\beta)] & 0 \\ \mu_2[(1+\beta) - (1-\beta)\kappa_1] & 0 & 0 & -\frac{1}{2}[\kappa_1(1-\beta) + \kappa_2(1+\beta)] \\ 0 & -\mu[(1+\beta) - (1-\beta)\kappa_1] & -\frac{1}{2}[\kappa_1(1-\beta) + \kappa_2(1+\beta)] & 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 0 & -\mu_1\beta & \frac{1}{2}[\kappa_1(1-\beta) + (1+\beta)] & 0 \\ -\mu_1\beta & 0 & 0 & \frac{1}{4}[\kappa_1(1-\beta) + (1+\beta)] \\ 0 & \mu_2[(1+\beta) + \kappa_1(1-\beta)] & \frac{1}{2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)] & 0 \\ \mu_2[(1+\beta) + \kappa_1(1-\beta)] & 0 & 0 & -\frac{1}{2}[\kappa_2(1+\beta) - \kappa_1(1-\beta)] \end{bmatrix}.$$

The system (9) is solved by diagonalization, as in Ni and Nemat-Nasser (1991, 1992), of the integral equations, and (9) is rewritten in the form

$$\frac{1}{\pi i}\left(\frac{1}{y} * \hat{p}\right) + (-iH_1^{-1}H_2)\hat{p} + iH_1^{-1}\hat{T} = 0 \quad (10)$$

with

$$H_1^{-1}H_2 = \begin{bmatrix} 0 & \frac{1-\kappa_2}{1+\kappa_2} & \frac{\kappa_1}{\mu(1+\kappa_1)} & 0 \\ -\frac{1-\kappa_2}{1+\kappa_2} & 0 & 0 & -\frac{\kappa_1}{\mu_1(1+\kappa_1)} \\ -\frac{4\mu_2}{1+\kappa_2} & 0 & 0 & \frac{1-\kappa_1}{1+\kappa_1} \\ 0 & -\frac{4\mu_2}{1+\kappa_2} & -\frac{1-\kappa_1}{1+\kappa_1} & 0 \end{bmatrix}. \quad (11)$$

2 Solution of the Governing Equations

The solution of (10) involves the eigenvalues s of $-iH_1^{-1}H_2$

$$|H_1^{-1}H_2 - isI| = 0$$

which, considering (11), are

$$s_j = \frac{1}{(1+\kappa_1)(1+\kappa_2)}\left[\pm(1-\kappa_1\kappa_2) \pm \sqrt{(\kappa_1-\kappa_2)^2 - \frac{4\mu_2\kappa_1}{\mu_1}(1+\kappa_1)(1+\kappa_2)}\right]_{j=1,2,3,4} \quad (12)$$

with

$$s_1 = \{+,+\},\ s_2 = \{+,-\},\ s_3 = \{-,+\},\ s_4 = \{-,-\}$$

so that $s_3 = -s_2$ and $s_4 = -s_1$.

The eigenvectors corresponding to s_1, s_2, s_3, s_4, respectively, are

$$\mathbf{v}_1 = \begin{pmatrix} \kappa_1(1+\kappa_2) \\ i\kappa_1(1+\kappa_2) \\ i\mu_1(1+\kappa_1)[(1-\kappa_2) - s_1(1+\kappa_2)] \\ -\mu_1(1+\kappa_1)[(1-\kappa_2) - s_1(1+\kappa_2)] \end{pmatrix}$$

$$\mathbf{v}_2 = \begin{pmatrix} \kappa_1(1+\kappa_2) \\ i\kappa_1(1+\kappa_2) \\ i\mu_1(1+\kappa_1)[(1-\kappa_2) - s_2(1+\kappa_2)] \\ -\mu_1(1+\kappa_1)[(1-\kappa_2) - s_2(1+\kappa_2)] \end{pmatrix}$$

$$\mathbf{v}_3 = \begin{pmatrix} \kappa_1(1+\kappa_2) \\ -i\kappa_1(1+\kappa_2) \\ -i\mu_1(1+\kappa_1)[(1-\kappa_2) + s_3(1+\kappa_2)] \\ -\mu_1(1+\kappa_1)[(1-\kappa_2) + s_3(1+\kappa_2)] \end{pmatrix}$$

$$\mathbf{v}_4 = \begin{pmatrix} \kappa_1(1+\kappa_2) \\ -i\kappa_1(1+\kappa_2) \\ -i\mu_1(1+\kappa_1)[(1-\kappa_2) + s_4(1+\kappa_2)] \\ -\mu_1(1+\kappa_1)[(1-\kappa_2) + s_4(1+\kappa_2)] \end{pmatrix}. \quad (13)$$

In order to proceed to the diagonalization of (10) we premultiply it by E^{-1}, where the matrix E is defined as $E = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]$, so that (10) yields

$$\frac{1}{\pi i}\frac{1}{y} * (E^{-1}\hat{p}) + E^{-1}(-iH_1^{-1}H_2)EE^{-1}\hat{p} = -iE^{-1}H_1^{-1}\hat{T}, \quad (14)$$

which, with the definitions $Q = E^{-1}\hat{p}$, $F \equiv -iE^{-1}H_1^{-1}T$, and the relation

$$E^{-1}(-iH_1^{-1}H_2)E = \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & s_3 & \\ & & & s_4 \end{bmatrix},$$

takes the final diagonalized form

$$\frac{1}{\pi i}\frac{1}{y} * Q + \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & s_3 & \\ & & & s_4 \end{bmatrix} Q = F. \quad (15)$$

Equation (15) has the solution

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$Q_j = \frac{s_j}{s_j^2 - 1} F_j + \frac{1}{\pi i (s_j^2 - 1)} \left(\frac{y + a}{a - y}\right)^{m_j} \times \int_{-a}^{a} \left(\frac{a - \zeta}{a + \zeta}\right)^{m_j} \frac{F_j d\zeta}{\zeta - y} + \frac{c_j}{(y + a)^{1 - m_j}(a - y)^{m_j}}, \quad j = 1, 2, 3, 4 \quad (16)$$

with

$$m_j = \frac{1}{2\pi i} \ln\left(\frac{s_j - 1}{s_j + 1}\right), \quad 0 < \mathrm{Re}(m_j) < 1. \quad (17)$$

The constants c_j are determined from the side conditions

$$\int_{-a}^{a} b_x(\xi) d\xi = \int_{-a}^{a} b_y(\xi) d\xi = \int_{-a}^{a} p_x(\xi) d\xi = \int_{-a}^{a} p_y(\xi) d\xi = 0, \quad (18)$$

expressing the fact that the total dislocation density vanishes and that the anticrack is not loaded by a concentrated force.

Application of (18) yields

$$c_j = \frac{2iam_j F_j}{(s_j^2 - 1) \sin(m_j \pi)}.$$

The rotation of the anticrack ω is determined by the condition that the moment of forces applied on the anticrack vanishes, i.e.,

$$\int_{-a}^{a} y p_x(\xi) d\xi = 0,$$

which yields

$$\hat{\omega} = \frac{-\hat{T}_{xy}\{a_3\mu_2[\kappa_1(1 - \beta) - (1 + \beta)] - a_2[\kappa_2(1 + \beta) + \kappa_1(1 - \beta)]\}}{\left\{\frac{a_2}{4}[\kappa_2(1 - \beta) - (1 + \beta)] - a_3\mu_1\right\}} \quad (19)$$

with

$$a_2 = \frac{\mu_1}{\kappa_1}[(1 - \kappa_2) - (1 + \kappa_2)s_2][(1 - \kappa_2) - (1 + \kappa_2)s_1][m_2(1 - m_2) - m_1(1 - m_1)]$$

$$a_3 = m_1(1 - m_1)[(1 - \kappa_2) - (1 + \kappa_2)s_1] - m_2(1 - m_2)[(1 - \kappa_2) - (1 + \kappa_2)s_1].$$

Finally, the solution for the dislocation and line load densities $\hat{p} = \begin{pmatrix} b_x \\ b_y \\ p_x \\ p_y \end{pmatrix}$ is

$$\hat{p} = E \begin{bmatrix} g(s_1) & & & \\ & g(s_2) & & \\ & & g(s_3) & \\ & & & g(s_4) \end{bmatrix} E^{-1} H_1^{-1} \hat{T} \quad (20)$$

with

$$g(s_j) = \frac{i}{(s_j^2 - 1) \sin m_j \pi} \left[\frac{y + a}{a - y}\right]^{m_j} \left[1 - \frac{2am_j}{y + a}\right]$$

$$H_1^{-1} = \frac{4}{m_2(1 - \beta^2)(1 + \kappa_1)(1 + \kappa_2)} \cdot \begin{bmatrix} -\frac{1}{2}[\kappa_2(1 + \beta) + \kappa_1(1 - \beta)] & 0 & -\frac{1}{4}[\kappa_1(1 - \beta) - (1 + \beta)] & 0 \\ 0 & -\frac{1}{2}[\kappa_2(1 + \beta) + \kappa_1(1 - \beta)] & 0 & \frac{1}{4}[\kappa_1(1 - \beta) - (1 + \beta)] \\ 0 & \mu_2[\kappa_1(1 - \beta) - (1 + \beta)] & 0 & -\mu_1 \\ \mu_2[\kappa_1(1 - \beta) - (1 + \beta)] & 0 & -\mu_1 & 0 \end{bmatrix} \quad (21)$$

3 Analysis of the Singularities at the Tip

From the form of the solution given by (20) and (21) the singularities obviously depend on the values of m_j given by (17), which in turn depend on the nature of the eigenvalues given by (12). The eigenvalues are real or complex according to the sign of the discriminant

$$\Delta^2 \equiv (\kappa_1 - \kappa_2)^2 - \frac{4\mu_2\kappa_1}{\mu_1}(1 + \kappa_1)(1 + \kappa_2). \quad (22)$$

The two cases are analyzed below:

Case i: $\Delta^2 \geq 0$. In this case s_j are real and we prove that

$$|s_j| \leq 1 \quad j = 1, 2, 3, 4 \quad (23)$$

by easily showing that

$$1 > \frac{(1 - \kappa_1\kappa_2) + |\kappa_1 - \kappa_2|}{(1 + \kappa_1)(1 + \kappa_2)} \geq s_1 \geq \frac{(1 - \kappa_1\kappa_2) + |\kappa_1 - \kappa_2|}{(1 + \kappa_1)(1 + \kappa_2)} > 0,$$

and similarily for s_2, s_3, s_4.

These make the argument of the logarithm in (17) real, and we can write

$$m_j = \frac{1}{2\pi i} \ln\left(\frac{s_j - 1}{s_j + 1}\right) = \frac{1}{2} + \frac{1}{2\pi i} \ln\left|\frac{s_j - 1}{s_j + 1}\right|. \quad (24)$$

In addition, by considering that $0 < \mathrm{Re}(m_j) < 1$, and that $s_1 = -s_4$, $s_2 = -s_3$, we have

Eigenvalue s	s_1	s_2	s_3	s_4
Singularity m	$\frac{1}{2} + i\gamma_1$	$\frac{1}{2} + i\gamma_2$	$\frac{1}{2} - i\gamma_2$	$\frac{1}{2} - i\gamma_1$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with

$$\gamma_j = -\frac{1}{2\pi}\ln\left|\frac{s_j - 1}{s_j + 1}\right|, \quad j = 1, 2.$$

Thus the singularities of (16) are square root with oscillatory component, with the singular term being

$$(a - y)^{-(1/2)}(a + y)^{-(1/2)}\left(\frac{a - y}{a + y}\right)^{\pm i\gamma_j}. \tag{25}$$

Case ii: $\Delta^2 < 0$. The four roots s_1, s_2, s_3, s_4 are complex:

$$s_1 = \frac{1}{(1 + \kappa_1)(1 + \kappa_2)}\{(1 - \kappa_1\kappa_2) + i|\Delta|\},$$

$$s_2 = \frac{1}{(1 + \kappa_1)(1 + \kappa_2)}\{(1 - \kappa_1\kappa_2) - i|\Delta|\}$$

$$s_3 = \frac{1}{(1 + \kappa_1)(1 + \kappa_2)}\{-(1 - \kappa_1\kappa_2) + i|\Delta|\},$$

$$s_4 = \frac{1}{(1 + \kappa_1)(1 + \kappa_2)}\{-(1 - \kappa_1\kappa_2) - i|\Delta|\}$$

so that in (17) the logarithm of a complex number has to be considered.

For s_1 we have

$$\ln\left(\frac{s_1 - 1}{1 + s_1}\right) = \ln\left[\frac{(AC + B^2) + iB(C - A)}{C^2 + B^2}\right] \tag{26}$$

with

$$A = -(2\kappa_1\kappa_2 + \kappa_1 + \kappa_2), \quad C = 2 + \kappa_1 + \kappa_2, \quad B = |\Delta|,$$

$$AC + B^2 = (1 + \kappa_1)(1 + \kappa_2)\left[\frac{4\mu_2\kappa_1}{\mu_1} - 2(\kappa_1 + \kappa_2)\right].$$

The location of the principal value of (26) will depend on the sign of the real part over the imaginary part; i.e., here on the sign of $AC + B^2$, which again, branches into two subcases:

(a) $G \equiv \dfrac{4\mu_2\kappa_1}{\mu_1} - 2(\kappa_1 + \kappa_2) \geq 0$ which implies

$$0 \leq \arg\left(\frac{s_1 - 1}{s_1 + 1}\right) \leq \frac{\pi}{2}$$

(b) $G \equiv \dfrac{4\mu_2\kappa_1}{\mu_1} - 2(\kappa_1 + \kappa_2) < 0$ which implies

$$\frac{\pi}{2} < \arg\left(\frac{s_1 - 1}{s_1 + 1}\right) < \pi. \tag{27}$$

With the notation

$$\theta = \arg\left(\frac{s_1 - 1}{s_1 + 1}\right), \quad \theta_1 = \frac{\theta}{2\pi},$$

case (a) implies:

$$0 \leq \theta \leq \frac{\pi}{2} \rightarrow 0 \leq \theta_1 \leq \frac{1}{4}, \quad \text{and} \quad (1 - \theta_1) \geq \frac{3}{4}$$

and case (b) implies:

$$\frac{\pi}{2} \leq \theta \leq \pi \rightarrow \frac{1}{4} < \theta_1 < \frac{1}{2}, \quad \text{and} \quad (1 - \theta_1) \leq \frac{3}{4}$$

with the exponent m_1 in (16) being $m_1 = \theta_1 + i\gamma$, with γ as defined in (28).

Analogously treating s_2 we have

$$\ln\left(\frac{s_2 - 1}{s_2 + 1}\right) = \ln\left[\frac{AC + B^2 - i(C - A)B}{C^2 + B^2}\right]$$

$$m_2 = \theta_2 + i\gamma, \quad \theta_2 = \frac{\theta}{2\pi}, \quad \theta = \arg\left(\frac{s_2 - 1}{s_2 + 1}\right).$$

case (a)

$$G \geq 0 \rightarrow \frac{3\pi}{2} \leq \theta \leq 2\pi \rightarrow \frac{3}{4} \leq \theta_2 \leq 1, \quad \theta_2 = 1 - \theta_1$$

case (b)

$$G \leq 0 \rightarrow \pi \leq \theta \leq \frac{3\pi}{2} \rightarrow \frac{1}{2} < \theta_2 \leq \frac{3}{4}, \quad \theta_2 = 1 - \theta_1$$

Proceeding similarly with s_3 and s_4, we can write the solution as follows:

	s_1	s_2	s_3	s_4
m	$\theta_1 + i\gamma$	$(1 - \theta_1) + i\gamma$	$\theta_1 - i\gamma$	$(1 - \theta_1) - i\gamma$

where

$$\theta_1 = \frac{1}{2\pi} \times \arctan\left[\frac{2\mu_2\kappa_1 - \mu_1(\kappa_1 + \kappa_2)}{\sqrt{4\mu_1\mu_2\kappa_1(1 + \kappa_1)(1 + \kappa_2) - \mu_1^2(\kappa_1 - \kappa_2)^2}}\right],$$

$$-\frac{\pi}{2} \leq \theta_1 \leq \frac{\pi}{2}$$

and

$$\gamma = \frac{1}{4\pi}\ln\left(\frac{\mu_1 + \mu_2\kappa_1}{\kappa_1(\mu_2 + \mu_1\kappa_2)}\right), \tag{28}$$

so that the singular terms of the solution are of the form

$$A_1(a - y)^{-\theta_1}(a + y)^{-(1-\theta_1)}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma} + A_2(a - y)^{-(1-\theta_1)}(a + y)^{-\theta_1}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma}. \tag{29}$$

A possibility of double root can arise when both materials are incompressible in plane strain ($\kappa_1 =$, $\kappa_2 = 1$) in which case the solution is of the form

$$A_1(a - y)^{-\theta_1}(a + y)^{-(1-\theta_1)}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma} + A_2(a - y)^{-\theta_1}(a + y)^{-(1-\theta_1)}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma}\ln\left(\frac{a + y}{a - y}\right) + A_3(a - y)^{-(1-\theta_1)}(a + y)^{-\theta_1}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma} + A_4(a - y)^{-(1-\theta_1)}(a + y)^{-\theta_1}\left(\frac{a + y}{a - y}\right)^{\pm i\gamma}\ln\left(\frac{a + y}{a - y}\right)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with

$$\theta_1 = \frac{1}{2\pi} \arctan\left(\frac{\mu_2 - \mu_1}{2\sqrt{\mu_2\mu_1}}\right)$$

and

$$\frac{1}{2} < \delta < \frac{3}{4} \quad \text{for} \quad \frac{\mu_2}{\mu_1} \leq 1$$

$$\frac{3}{4} \leq \delta < 1 \quad \text{for} \quad \frac{\mu_2}{\mu_1} \geq 1.$$

Another possibility of double root is when $\Delta = 0$ in (22), in which case the singularity is 1/2, and the solution is as in (25), with the additional term

$$(a - y)^{-(1/2)}(a + y)^{-(1/2)}\left(\frac{a - y}{a + y}\right)^{\pm i\gamma_j} \ln\left(\frac{a + y}{a - y}\right).$$

The dependence of the leading singular term $r^{-\delta}$ of the solution (29) in terms of μ_2/μ_1 is, according to the previous analysis, summarized as follows:

Ratio of Shear Moduli μ_2/μ_1	Singularity $r^{-\delta}$
$\frac{\mu_2}{\mu_1} \leq F(\kappa_1, \kappa_2)$	$\delta = \frac{1}{2}$
$F(\kappa_1, \kappa_2) \leq \frac{\mu_2}{\mu_1} \leq H(\kappa_1, \kappa_2)$	$\frac{1}{2} \leq \delta \leq \frac{1}{2}$
$\frac{\mu_2}{\mu_1} \geq H(\kappa_1, \kappa_2)$	$\frac{3}{4} \leq \delta \leq 1$

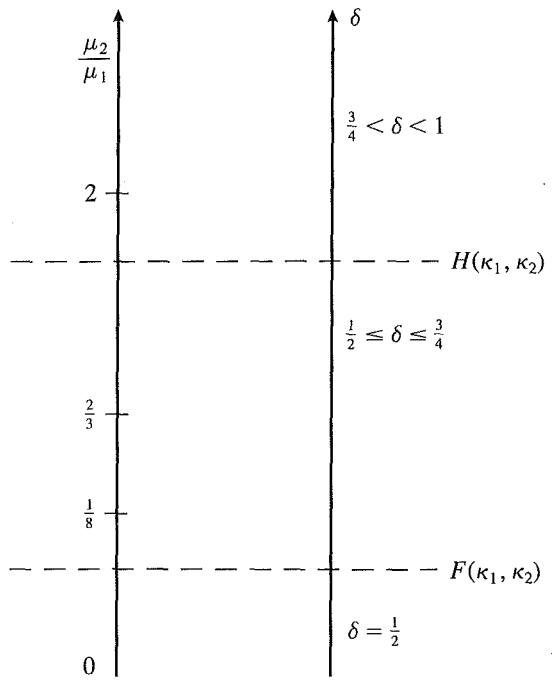
where

$$F(\kappa_1, \kappa_2) \equiv \frac{(\kappa_1 - \kappa_2)^2}{4\kappa_1(1 + \kappa_1)(1 + \kappa_2)}, \quad H(\kappa_1, \kappa_2) \equiv \frac{\kappa_1 + \kappa_2}{2\kappa_1}.$$

For the physical range of elastic constants $1 \leq \kappa_j \leq 3$, F and H vary in the range

$$0 \leq F(\kappa_1, \kappa_2) \leq \tfrac{1}{8}, \quad \tfrac{2}{3} \leq H(\kappa_1, \kappa_2) \leq 2$$

and the relations (30) are depicted on the graph:



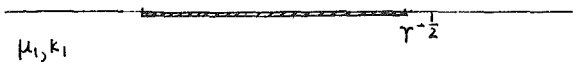


Fig. 2 Half-space with a bonded rigid line inclusion

4 Physical Interpretation of the Solution

In order to understand the physical behavior of the solution we look at the limiting cases of the following.

(1) The two materials being *identical,* i.e., $\mu_1 = \mu_2$, $\kappa_1 = \kappa_2$. In this case $\Delta^2 < 0$, $\theta_1 = (1/2\pi) \arctan(\infty) = (1/2\pi)(\pi/2) = 1/4$ and the leading singularity is $1 - (1/4) = 3/4$, consistent with the solution earlier obtained for a debonded anticrack in a uniform material by Gdoutos et al. (1989), and also with the singularity at the vertex of an infinite wedge of angle 2π which is stress free on one flank and clamped on the other (Williams, 1952).

(2) $\mu_2/\mu_1 \to 0$, i.e., either $\mu_2 \to 0$ or $\mu_1 \to \infty$, that is either the upper material vanishes, or the lower material is infinitely rigid. If the upper material vanishes then the problem reduces to a rectangular punch on a half-space. In this case $\Delta^2 > 0$ and the singularity is $\frac{1}{2}$, consistent with the singularity at the corner of a punch. (Fig. 2)

(3) $\mu_2/\mu_1 \to \infty$, with the upper material being infinitely rigid. In this case $\Delta^2 \to -\infty < 0$, $\theta_1 = 0$ and the singularity is $(a - y)^{-1}$, $(a + y)^{-1}$. This is equivalent to a dislocation at $y = \pm a$, since the only way for the rigid inclusion to move is as a rigid body with translation and rotation.

Equations (30) and the limiting cases can be summarized as follows: When the upper material is absent the behavior is one of a rigid punch in an elastic half-space, with the singularity being 1/2. The singularity remains 1/2 when the upper material is very soft, and when the ratio μ_2/μ_1 exceeds the value of $F(\kappa_1, \kappa_2)$ it starts increasing from 1/2 to 3/4, achieved when μ_2/μ_1 reaches the value $H(\kappa_1, \kappa_2)$. When the two materials are identical the singularity is 3/4. When μ_2/μ_1 increases further the singularity becomes bigger than 3/4 and tends to 1 as the upper material tends to becoming infinitely rigid.

In brief, we can say that while the problem of a single crack or a single anticrack at an interface induces square-root singular fields, the problem of the debonded anticrack entails a greater amount of incompatibility and hence greater singularities appear; that is more singular stresses are required from the material in order to accommodate the boundary conditions.

5 Interaction of the Debonded Anticrack With a Dislocation or Concentrated Force

It was shown previously (Markenscoff et al., 1994; Dundurs and Markenscoff, 1989), that the crack absorbs the dislocation (an r^{-1} singularity) as it approaches the crack-tip while it does not absorb the concentrated force (also an r^{-1} singularity). Dually, the anticrack absorbs the concentrated force, while it does not absorb the dislocation. As previously mentioned the method of solving crack and anticrack problems by distributing dislocation densities and line load densities, respectively, has a physical basis versus the equivalent method of complex variables, and readily explains the absorption of the dislocation/concentrated force by a crack/anticrack, respectively, since the crack/anticrack is nothing else but a distribution of dislocation forces.

The question arises then whether the debonded anticrack, which consists of both a distribution of dislocations and line loads absorbs the dislocation and the concentrated force. Although we have not yet done the analysis, we conjecture that both are absorbed. The conjecture is based on the physical argument that in the case of the upper material absent the anticrack absorbs the force, and adding the upper material will

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

only "make things better" for the deformation. Indeed, when a concentrated force approaches the tip of an anticrack the deformation under the concentrated force is distributed by the anticrack, which moves as a rigid body. This results in "the absorbion" of the force also confirmed by the solution of the Hertz rectangular rigid punch problem on a half-plane. With an upper material added, the deformation will be additionally reduced and the force will of course still be "absorbed."

Acknowledgment

This support of the National Science Foundation (XM) MSS-9313094 and INT-91-15553, and the Center of Excellence for Advanced Materials at the University of California, San Diego (L.N.) is gratefully acknowledged. Discussions with Prof. J. Dundurs are gratefully acknowledged.

References

Ballarini, R., 1990, "A Rigid Line Inclusion at a Bimaterial Interface," *Engineering Fracture Mechanics,* Vol. 37, pp. 173–182.

Dempsey, J. P., and Sinclair, G. B., 1981, "On the Singular Behavior at the Vertex of a Bi-Material Wedge," *Journal of Elasticity,* Vol. 11, No. 3, pp. 317–327.

Dundurs, J., 1969, discussion of a paper by D. B. Bogy, ASME JOURNAL OF APPLIED MECHANICS, pp. 650–652.

Dundurs, J., and Markenscoff, X., 1989, "A Green's Function Formulation of Anticracks and Their Interaction With Load-Induced Singularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 550–555.

Erdogan, F., 1969, "Approximate Solutions of Systems of Singular Integral Equations," *SIAM Journal of Applied Mathematics,* Vol. 17, pp. 1041–1059.

Erdogan, F., 1965, "Stress Distribution in Bonded Dissimilar Materials Containing Circular or Angular-Shaped Cavities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 829–836.

Gdoutos, E. G., Kourounis, C. G., Kattis, M. A., and Zacharopoulos, D. A., 1989, "A Partly Unbonded Rigid Fiber Inclusion in an Infinite Matrix," *Advances in Fracture Mechanics* K. Salamaa, Ravi-Chandar, Taplui, and Rao, eds., pp. 223–227.

Keer, L. M., 1975, "Mixed Boundary Value Problems for a Penny-Shaped Cut," *Journal of Elasticity,* Vol. 5, pp. 89–98.

Markenscoff, X., Ni, L., and Dundurs, J., 1994, "The Interface Anticrack: Solutions and Green's Functions for the Interaction of Cracks/Anticracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 61, pp. 797–802.

Muskhelishvili, N. I., 1953, *Some Basic Problems in the Theory of Elasticity,* Noorhoff, Leyden, The Netherlands.

Ni, L., and Nemat-Nasser, S., 1991, "Interface Cracks in Anisotropic Dissimilar Materials: An Analytic Solution," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, No. 1, pp. 113–144.

Ni, L., and Nemat-Nasser, S., 1992, "Interface Cracks in Anisotropic Dissimilar Materials: General Case," *Quarterly of Applied Mathematics,* Vol. L, No. 2, pp. 305–322.

Sherman, D. I., 1940, "The Mixed Problems of Potential Theory and of the Theory of Elasticity for the Plane With a Finite Number of Straight Cuts," *Comptes Rendus de L'Academie des Sciences de L'U.R.S.,* Vol. 27, pp. 330–334.

Ting, T. C. T., 1986, "Explicit Solution and Invariance of the Singularities at an Interface Crack in Anisotropic Composites," *International Journal of Solids and Structures,* Vol. 22, No. 9, pp. 965–983.

Wang, Z. Y., Zhang, H. T., and Chou, Y. T., 1985, "Characteristics of the Elastic Field of a Rigid Line in Homogeneity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 818–822.

Williams, M. L., 1952, "Stress Singularities Resulting From Various Boundary Conditions in Angular Corners of Plates in Extension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, pp. 526–528.

Wu, K. C., 1990, "Line Inclusions at Anisotropic Bimaterial Interface," *Mechanics and Materials,* Vol. 10, pp. 173–182.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

An Exact Three-Dimensional Solution for Simply Supported Rectangular Piezoelectric Plates

P. Bisegna

F. Maceri

Department of Civil Engineering,
University of Rome "Tor Vergata,"
00133 Rome, Italy

An exact three-dimensional solution for the problem of a simply supported rectangular homogeneous piezoelectric plate is obtained, in the framework of the linear theory of piezoelectricity. The plate is made of a transversely isotropic material, is earthed on the lateral boundary, and is subjected to prescribed surface charge and tractions on the end faces. The limit of this solution as the plate thickness aspect ratio approaches zero is explicitly carried out. The analytical results obtained may constitute a reference case when developing or applying two-dimensional plate theories for the analysis of more complex piezoelectric problems. A numerical investigation in the case of a square uniformly loaded plate is also performed, in order to evaluate the influence of the thickness-to-side ratio on the three-dimensional solution of the plate problem.

1 Introduction

At the beginning of the present century, piezoelectricity was applied to sonars, to frequency control, to acoustic interferometry (Mason, 1981). Presently, applications of piezoelectric materials are mainly oriented towards detecting deformations and motions of structures, as well as in active structural control. As a matter of fact, by bonding or embedding a piezoelectric element in a structure, it is possible both to measure strains and displacements, and to give localized strains through which the deformation of the structure can be controlled. In this way a so-called "smart" structure is obtained. An effective sensing and control of a smart structure can be achieved only through a careful modelization of the coupled electroelastic behavior of the structure. The literature on this topic is wide: extensive references on structural piezoelectric sensing and control can be found in a book by Tzou (1993).

There is a special interest in the structural problem of piezoelectric plates, since it arises in modeling flat piezoelectric sensors and actuators. In the analysis of this problem approximate theories are often adopted: indeed, simplifying a priori hypotheses are assumed, concerning the direction of the electric field, and the representation form of the displacement field. As a consequence, the interest for exact solutions (i.e., solutions obtained from the full three-dimensional theory) arises, for the purpose of verifying the accuracy of the results provided by approximate theories or computations.

In this paper the framework of the linear theory of piezoelectricity is adopted, and an exact three-dimensional solution for the problem of a simply supported rectangular homogeneous piezoelectric plate is obtained. The plate is made of a transversely isotropic material, is earthed on the lateral boundary, and is subjected to prescribed surface charge and tractions on the end faces. The given loads are represented in double-series trigonometric form. The expressions of all the mechanical and electric involved functions are explicitly found.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Jan. 23, 1995; final revision, Nov. 8, 1995. Associate Technical Editor: R. Abeyaratne.

The solving technique adopted here is based on the separation of the independent variables. It was applied to study plates regarded as three-dimensional bodies by some researchers. In particular, Vlasov (1957) and Levinson (1984) studied the problem of isotropic homogeneous elastic plates; Pan (1991) considered the problem of transversely isotropic homogeneous elastic plates; Lee (1967), Srinivas and Rao (1970), Pagano (1970), and Noor and Scott Burton (1990) studied the problem of orthotropic or anisotropic composite elastic plates. An application of the same technique to the problem of piezoelectric plates is due to Ray et al. (1992), which considered only the simplest case of cylindrical bending. In this paper the more general case of bidirectional bending is studied.

The paper is organized as follows. Section 2 is devoted to notation. The linear piezoelectric constitutive equations and the basic equations of the linear theory of piezoelectricty are recalled in Sections 3 and 4, respectively. The plate problem is introduced in Section 5 and a separate-variables three-dimensional solution is obtained in Section 6. This solution is then expanded into a power series of the thickness of the plate, and the lowest-order term of the expansion is explicitly carried out (Section 7). As a consequence, the limit of the three-dimensional solution of the plate problem as the plate thickness aspect ratio approaches zero is obtained, in a closed form. This analytical result is intended to serve as a reference case when developing or applying approximate theories for the analysis of more complex piezoelectric problems. Moreover, it may be useful in the development of a consistent theory of thin piezoelectric plates, since it supplies the orders of infinity of the mechanical and electric quantities as the plate thickness aspect ratio approaches zero. At authors' knowledge, an explicit calculation, for any load condition, of the limit of all the mechanical and electric quantities as the plate thickness aspect ratio approaches zero is carried out here for the first time. Levinson (1984) and Pan (1991) considered only the limit of the deflection of an elastic plate for one particular load condition.

Finally, a numerical investigation in the case of a square, uniformly loaded plate is performed in Section 8, in order to evaluate the influence of the thickness-to-side ratio on the three-dimensional solution of the plate problem.

2 Notation

Let $\mathcal{E}$ be the three-dimensional Euclidean space, and let $\mathcal{V}$ be the vector space associated to $\mathcal{E}$. Let $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ be an orthonormal basis of $\mathcal{V}$ and let a_i $(i = 1 \ldots 3)$ denote the components

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of a vector $\mathbf{a} \in \mathcal{V}$ with respect to this basis. Let $O \in \mathcal{E}$ be a fixed origin and let (x_1, x_2, x_3) denote the coordinates of the typical point of $\mathcal{E}$ in the cartesian frame $(O, \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$.

Let *Lin* be the vector space of the second-order tensors on $\mathcal{V}$, *Sym* $\subset$ *Lin* the subspace of the symmetric tensors, and *Orth* $\subset$ *Lin* the group of the isometries of $\mathcal{V}$. The identity of $\mathcal{V}$ is denoted by $\mathbf{I}$; the symmetric part of $\mathbf{A} \in Lin$ is denoted by sym $\mathbf{A}$. *Lin* is endowed with the usual inner product defined by the trace operator. The symbol "$\cdot$" denotes the scalar product both between vectors and between tensors. The symbol "$\otimes$" denotes the tensor product between vectors, between tensors, and between a vector and a tensor.

An orthogonal (not orthonormal) basis of *Sym* is $\{\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{H}_5, \mathbf{H}_6\}$, where

$$\mathbf{H}_1 := \mathbf{c}_1 \otimes \mathbf{c}_1, \quad \mathbf{H}_2 := \mathbf{c}_2 \otimes \mathbf{c}_2, \quad \mathbf{H}_3 := \mathbf{c}_3 \otimes \mathbf{c}_3,$$

$$\mathbf{H}_4 := \text{sym } \mathbf{c}_2 \otimes \mathbf{c}_3, \quad \mathbf{H}_5 := \text{sym } \mathbf{c}_3 \otimes \mathbf{c}_1,$$

$$\mathbf{H}_6 := \text{sym } \mathbf{c}_1 \otimes \mathbf{c}_2. \tag{1}$$

The components $\mathbf{A} \cdot \mathbf{H}_i$ of a tensor $\mathbf{A} \in Sym$ with respect to this basis are denoted by A_i $(i = 1 \ldots 6)$.

The symbols "∇" and "div" represent the gradient and the divergence operator, respectively.

3 Linear Piezoelectric Constitutive Equations

The framework of the infinitesimal strain theory is adopted. Let the strain tensor be $\mathbf{E} \in Sym$, the stress tensor be $\mathbf{T} \in Sym$, the electric field vector be $\mathbf{e} \in \mathcal{V}$ and the electric displacement vector be $\mathbf{d} \in \mathcal{V}$. The enthalpy $\mathcal{H}$ per unit volume of a piezoelectric material (Parton, 1984) can be expressed as a function of the strain $\mathbf{E}$ and of the electric field $\mathbf{e}$:

$$\mathcal{H} = \mathcal{H}(\mathbf{E}, \mathbf{e}). \tag{2}$$

The stress $\mathbf{T}$ and the electric displacement $\mathbf{d}$ are the dual quantities of $\mathbf{E}$ and $\mathbf{e}$, respectively:

$$\mathbf{T} = \frac{\partial \mathcal{H}}{\partial \mathbf{E}}, \quad \mathbf{d} = -\frac{\partial \mathcal{H}}{\partial \mathbf{e}}. \tag{3}$$

By definition, for a *linear* piezoelectric material the enthalpy $\mathcal{H}$ is a quadratic homogeneous function of $\mathbf{E}$ and $\mathbf{e}$: hence, it takes the form

$$\mathcal{H}(\mathbf{E}, \mathbf{e}) = \tfrac{1}{2}\mathcal{C}^{\mathbf{e}}[\mathbf{E}] \cdot \mathbf{E} - L[\mathbf{E}] \cdot \mathbf{e} - \tfrac{1}{2}\epsilon^{\mathbf{E}}[\mathbf{e}] \cdot \mathbf{e}, \tag{4}$$

where $\mathcal{C}^{\mathbf{e}}$ is a fourth-order symmetric tensor (elasticity tensor, evaluated at constant electric field), $\epsilon^{\mathbf{E}}$ is a second-order symmetric tensor (dielectric tensor, evaluated at constant strain) and L is a third-order tensor (piezoelectric tensor). As a consequence, the constitutive equations of a linear piezoelectric material can be formulated as

$$\begin{cases} \mathbf{T}(\mathbf{E}, \mathbf{e}) = \mathcal{C}^{\mathbf{e}}[\mathbf{E}] - L^t[\mathbf{e}] \\ \mathbf{d}(\mathbf{E}, \mathbf{e}) = L[\mathbf{E}] + \epsilon^{\mathbf{E}}[\mathbf{e}], \end{cases} \tag{5}$$

where L^t denotes the transpose of L, defined by

$$\mathbf{A} \cdot L^t[\mathbf{b}] = L[\mathbf{A}] \cdot \mathbf{b}, \quad \mathbf{A} \in Lin, \quad \mathbf{b} \in \mathcal{V}. \tag{6}$$

The tensors $\mathcal{C}^{\mathbf{e}}$, L, and $\epsilon^{\mathbf{E}}$ depend, respectively, upon 21, 18, and 6 material constants (which are called elastic, piezoelectric, and dielectric constants, respectively). The number of a priori different material constants is reduced by taking into account the symmetries of the material. The symmetry group of a piezoelectric material is defined as (Landau and Lifshitz, 1981):

$$\mathcal{G} := \{\mathbf{Q} \in Orth: \mathbf{QT}(\mathbf{E}, \mathbf{e})\mathbf{Q}^t = \mathbf{T}(\mathbf{QEQ}^t, \mathbf{Qe}); \quad \mathbf{Qd}(\mathbf{E}, \mathbf{e}) = \mathbf{d}(\mathbf{QEQ}^t, \mathbf{Qe})\}. \tag{7}$$

In this paper, transversely isotropic piezoelectric materials (class ∞ mm), having the transverse isotropy axis parallel to $\mathbf{c}_3$, are considered: their symmetry group contains the rotations around $\mathbf{c}_3$ and the reflections with respect to planes parallel to $\mathbf{c}_3$ (Ikeda, 1990). The representation formulas of $\mathcal{C}^{\mathbf{e}}$, L, and $\epsilon^{\mathbf{E}}$ for these materials are (Varadan et al., 1987)

$$\begin{cases} \mathcal{C}^{\mathbf{e}} = C_{11}^{\mathbf{e}}(\mathbf{H}_1 \otimes \mathbf{H}_1 + \mathbf{H}_2 \otimes \mathbf{H}_2) + C_{12}^{\mathbf{e}}(\mathbf{H}_1 \otimes \mathbf{H}_2 + \mathbf{H}_2 \otimes \mathbf{H}_1) + C_{13}^{\mathbf{e}}(\mathbf{H}_1 \otimes \mathbf{H}_3 + \mathbf{H}_3 \otimes \mathbf{H}_1 + \mathbf{H}_2 \otimes \mathbf{H}_3 + \mathbf{H}_3 \otimes \mathbf{H}_2) + C_{33}^{\mathbf{e}}\mathbf{H}_3 \otimes \mathbf{H}_3 + 4C_{44}^{\mathbf{e}}(\mathbf{H}_4 \otimes \mathbf{H}_4 + \mathbf{H}_5 \otimes \mathbf{H}_5) + 4C_{66}^{\mathbf{e}}\mathbf{H}_6 \otimes \mathbf{H}_6 \\ L = L_{31}(\mathbf{c}_3 \otimes \mathbf{H}_1 + \mathbf{c}_3 \otimes \mathbf{H}_2) + L_{33}\mathbf{c}_3 \otimes \mathbf{H}_3 + 2L_{15}(\mathbf{c}_1 \otimes \mathbf{H}_5 + \mathbf{c}_2 \otimes \mathbf{H}_4) \\ \epsilon^{\mathbf{E}} = \epsilon_{11}^{\mathbf{E}}(\mathbf{H}_1 + \mathbf{H}_2) + \epsilon_{33}^{\mathbf{E}}\mathbf{H}_3, \end{cases} \tag{8}$$

where $C_{66}^{\mathbf{e}} := (C_{11}^{\mathbf{e}} - C_{12}^{\mathbf{e}})/2$ and $C_{11}^{\mathbf{e}}$, $C_{12}^{\mathbf{e}}$, $C_{13}^{\mathbf{e}}$, $C_{33}^{\mathbf{e}}$, $C_{44}^{\mathbf{e}}$, L_{31}, L_{33}, L_{15}, $\epsilon_{11}^{\mathbf{E}}$, $\epsilon_{33}^{\mathbf{E}}$ are ten independent material constants.

The representation formulas (8) hold also for hexagonal piezoelectric materials belonging to the class C_{6v} (or 6 mm) with the sixth-order symmetry axis parallel to $\mathbf{c}_3$ (Kiral and Eringen, 1990). As a consequence, the results obtained here hold also for the materials of this class.

In the technical literature one finds more frequently Eqs. (5) expressed in matrix notation. The components of $\mathbf{E}$, $\mathbf{T}$, $\mathbf{e}$, $\mathbf{d}$ are arranged in form of vectors, while the material constants are arranged in a square matrix. When the representation formulas (8) hold, it turns out

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} C_{11}^{\mathbf{e}} & C_{12}^{\mathbf{e}} & C_{13}^{\mathbf{e}} & 0 & 0 & 0 & 0 & 0 & -L_{31} \\ C_{12}^{\mathbf{e}} & C_{11}^{\mathbf{e}} & C_{13}^{\mathbf{e}} & 0 & 0 & 0 & 0 & 0 & -L_{31} \\ C_{13}^{\mathbf{e}} & C_{13}^{\mathbf{e}} & C_{33}^{\mathbf{e}} & 0 & 0 & 0 & 0 & 0 & -L_{33} \\ 0 & 0 & 0 & C_{44}^{\mathbf{e}} & 0 & 0 & 0 & -L_{15} & 0 \\ 0 & 0 & 0 & 0 & C_{44}^{\mathbf{e}} & 0 & -L_{15} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66}^{\mathbf{e}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_{15} & 0 & \epsilon_{11}^{\mathbf{E}} & 0 & 0 \\ 0 & 0 & 0 & L_{15} & 0 & 0 & 0 & \epsilon_{11}^{\mathbf{E}} & 0 \\ L_{31} & L_{31} & L_{33} & 0 & 0 & 0 & 0 & 0 & \epsilon_{33}^{\mathbf{E}} \end{bmatrix} \cdot \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ 2E_4 \\ 2E_5 \\ 2E_6 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix}. \tag{9}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Elastic, piezoelectric, and dielectric constants of some materials

	CdS	PZT-4	PZT-5H	PZT-8	C-24
$C^{\mathbf{e}}_{11}$ [GPa]	90.7	139.	126.	137.	150.
$C^{\mathbf{e}}_{33}$ [GPa]	93.8	115.	117.	124.	128.
$C^{\mathbf{e}}_{12}$ [GPa]	58.1	77.8	79.5	69.7	36.8
$C^{\mathbf{e}}_{13}$ [GPa]	50.9	74.3	84.1	71.6	30.9
$C^{\mathbf{e}}_{44}$ [GPa]	15.0	25.6	23.0	31.4	55.2
L_{31} [C/m^2]	−0.245	−5.2	−6.5	−4.0	1.51
L_{33} [C/m^2]	0.440	15.1	23.3	13.8	8.53
L_{15} [C/m^2]	−0.210	12.7	17.0	10.4	3.89
$\epsilon^{\mathbf{E}}_{11}$ [nF/m]	79.8	6.46	15.0	7.95	1.82
$\epsilon^{\mathbf{E}}_{33}$ [nF/m]	84.3	5.62	13.0	5.15	1.28

The material constants (at room temperature) of some piezoelectric materials are given in Table 1 (Eringen and Maugin, 1990; Smith, 1992; Toshiba, 1983; Vernitron, 1983): a 6 mm hexagonal crystal (Cadmium sulphure) and four transversely isotropic ceramics are considered. The units of the International System are used: gigaPascal for elastic constants, nanoFarad per meter for dielectric constants, and Coulomb per square meter for piezoelectric constants.

4 Basic Equations of the Linear Theory of Piezoelectricity

Let a body Ω be given, subjected to

(i) volume forces $\mathbf{b}$;
(ii) volume charge ρ;
(iii) surface forces $\mathbf{p}$ on a part $\partial_f\Omega$ of the boundary $\partial\Omega$;
(iv) surface charge ω on a part $\partial_\omega\Omega$ of $\partial\Omega$;
(v) prescribed displacement $\mathbf{s}_0$ on $\partial_s\Omega := \partial\Omega\backslash\partial_f\Omega$;
(vi) prescribed electric potential ϕ_0 on $\partial_\phi\Omega := \partial\Omega\backslash\partial_\omega\Omega$.

The strain field $\mathbf{E}$ corresponding to a displacement field $\mathbf{s}$ from the reference configuration of Ω to a deformed configuration is defined by the equation

$$\mathbf{E} = \text{sym } \nabla\mathbf{s}. \tag{10}$$

In this paper a time-independent (i.e., electrostatic) problem is considered: hence, the electric field can be derived from a potential, say ϕ:

$$\mathbf{e} = -\nabla\phi. \tag{11}$$

A displacement field $\mathbf{s}$ in Ω which equals $\mathbf{s}_0$ on $\partial_s\Omega$ and an electric potential field ϕ in Ω which equals ϕ_0 on $\partial_\phi\Omega$ are sought for, such that the stress field $\mathbf{T}$ and the electric displacement field $\mathbf{d}$, corresponding through Eqs. (5) to the strain field (10) and to the electric field (11), verify the balance equations (Maugin, 1988):

$$\begin{cases} \text{div } \mathbf{T} + \mathbf{b} = 0 & \text{in } \Omega \\ \text{div } \mathbf{d} - \rho = 0 & \text{in } \Omega \end{cases}$$

$$\text{and } \begin{cases} \mathbf{Tn} = \mathbf{p} & \text{on } \partial_f\Omega, \\ \{\mathbf{d}\}\cdot\mathbf{n} = \omega & \text{on } \partial_\omega\Omega, \end{cases} \tag{12}$$

where $\mathbf{n}$ is the outward normal field to $\partial\Omega$ and $\{\mathbf{d}\}$ denotes the jump of $\mathbf{d}$ across $\partial\Omega$.

It is assumed that outside Ω the electric displacement field vanishes: e.g., this assumption holds true, if Ω is surrounded by conductors, and is usually acceptable, if Ω is surrounded by the vacuum (or air). Hence, the equation $\{\mathbf{d}\}\cdot\mathbf{n} = \omega$ becomes

$$\mathbf{d}\cdot\mathbf{n} = -\omega \quad \text{on } \partial_\omega\Omega, \tag{13}$$

where $\mathbf{d}$ is the electric displacement field inside Ω.

As a consequence, the unknowns $\mathbf{s}$ and ϕ must satisfy the field equations

$$\begin{cases} \text{div } (\mathcal{C}^{\mathbf{e}}[\text{sym } \nabla\mathbf{s}]) + \text{div } (L^t[\nabla\phi]) + \mathbf{b} = 0 \\ \text{div } (L[\text{sym } \nabla\mathbf{s}]) - \text{div } (\epsilon^{\mathbf{E}}[\nabla\phi]) - \rho = 0 \end{cases} \text{in } \Omega, \tag{14}$$

and the boundary conditions

$$\begin{cases} \mathcal{C}^{\mathbf{e}}[\text{sym } \nabla\mathbf{s}]\mathbf{n} + L^t[\nabla\phi]\mathbf{n} = \mathbf{p} & \text{on } \partial_f\Omega \\ L[\text{sym } \nabla\mathbf{s}]\cdot\mathbf{n} - \epsilon^{\mathbf{E}}[\nabla\phi]\cdot\mathbf{n} = -\omega & \text{on } \partial_\omega\Omega \\ \mathbf{s} = \mathbf{s}_0 & \text{on } \partial_s\Omega \\ \phi = \phi_0 & \text{on } \partial_\phi\Omega. \end{cases} \tag{15}$$

5 The Simply Supported Rectangular Plate Problem

A rectangular plate having in-plane dimensions L_1 and L_2 and thickness H is considered. Let the reference configuration of the plate be the region

$$\Omega = (-L_1/2, L_1/2) \times (-L_2/2, L_2/2) \times (-H/2, H/2) \subset \mathcal{E}. \tag{16}$$

The end faces of the plate, orthogonal to the x_3-axis at $x_3 = \pm H/2$, are denoted as upper or lower face, while the other faces constitute the lateral boundary.

The plate is assumed to be made of a transversely isotropic material, with the transverse-isotropy axis parallel to the x_3-axis.

The volume forces $\mathbf{b}$ and the volume charge ρ are assumed to vanish. The boundary conditions on $\partial\Omega$ are assigned in a slightly more general way than that specified in the previous section. Indeed, a so-called "general boundary value problem," or "mixed-mixed problem" (Gurtin, 1972) is considered here, in order to model a simply supported plate. In particular, on the lateral boundary of the plate, the tangential component $(\mathbf{I} - \mathbf{n} \otimes \mathbf{n})\mathbf{s}_0$ of the displacement and the normal component $(\mathbf{n} \otimes \mathbf{n})\mathbf{p}$ of the surface forces are assumed to vanish (see, e.g., Pagano, 1970; Srinivas and Rao, 1970; Levinson, 1984; Pan, 1991). In addition, the potential ϕ is assumed to be zero. On the upper and lower face the surface forces and the surface charge are given. In formulas, the boundary conditions are as follows:

(i) on the lateral faces orthogonal to the x_1 and x_2-axis, respectively:

$$\begin{cases} s_2(\pm L_1/2, x_2, x_3) = 0 \\ s_3(\pm L_1/2, x_2, x_3) = 0 \\ \phi(\pm L_1/2, x_2, x_3) = 0 \\ T_1(\pm L_1/2, x_2, x_3) = 0 \end{cases} \text{and} \quad \begin{cases} s_1(x_1, \pm L_2/2, x_3) = 0 \\ s_3(x_1, \pm L_2/2, x_3) = 0 \\ \phi(x_1, \pm L_2/2, x_3) = 0 \\ T_2(x_1, \pm L_2/2, x_3) = 0; \end{cases} \tag{17}$$

(ii) on the upper and lower face:

$$\begin{cases} T_5(x_1, x_2, \pm H/2) = \pm p_1^{\pm}(x_1, x_2) \\ T_4(x_1, x_2, \pm H/2) = \pm p_2^{\pm}(x_1, x_2) \\ T_3(x_1, x_2, \pm H/2) = \pm p_3^{\pm}(x_1, x_2) \\ d_3(x_1, x_2, \pm H/2) = \mp\omega^{\pm}(x_1, x_2), \end{cases} \tag{18}$$

where $p_1^{\pm}$, $p_2^{\pm}$, and $p_3^{\pm}$ are the components of the surface forces $\mathbf{p}^{\pm}$, $\omega^{\pm}$ is the surface charge, and upper (lower) sign applies to upper (lower) face. Equations (17) and (18) are intended to

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

be read twice: one time taking all the upper signs, the other time taking all the lower signs.

The load acting on the plate is assumed to be symmetric with respect to the plane containing the x_1 and x_3-axes, and to the plane containing the x_2 and x_3-axes. In other words, the functions $p_3^{\pm}$ and $\omega^{\pm}$ are assumed to be even with respect to both x_1 and x_2; the functions $p_1^{\pm}$ are assumed to be odd with respect to x_1 and even with respect to x_2; the functions $p_2^{\pm}$ are assumed to be even with respect to x_1 and odd with respect to x_2.

With a view toward satisfying the boundary conditions (18), the functions $p_1^{\pm}$, $p_2^{\pm}$, $p_3^{\pm}$ and $\omega^{\pm}$ are expanded into a double Fourier series in the rectangle $(-L_1/2, L_1/2) \times (-L_2/2, L_2/2)$. As a consequence, it is sufficient to solve the problem of a plate subjected to surface forces and surface charge on the upper and lower face of the form

$$\begin{cases} p_1^{\pm}(x_1, x_2) = p_{1,n_1,n_2}^{\pm} \sin \dfrac{(2n_1+1)\pi x_1}{L_1} \cos \dfrac{(2n_2+1)\pi x_2}{L_2} \\ p_2^{\pm}(x_1, x_2) = p_{2,n_1,n_2}^{\pm} \cos \dfrac{(2n_1+1)\pi x_1}{L_1} \sin \dfrac{(2n_2+1)\pi x_2}{L_2} \\ p_3^{\pm}(x_1, x_2) = p_{3,n_1,n_2}^{\pm} \cos \dfrac{(2n_1+1)\pi x_1}{L_1} \cos \dfrac{(2n_2+1)\pi x_2}{L_2} \\ \omega^{\pm}(x_1, x_2) = \omega_{n_1,n_2}^{\pm} \cos \dfrac{(2n_1+1)\pi x_1}{L_1} \cos \dfrac{(2n_2+1)\pi x_2}{L_2}, \end{cases} \tag{19}$$

where n_1 and n_2 are fixed non-negative integers and $p_{1,n_1,n_2}^{\pm}$, $p_{2,n_1,n_2}^{\pm}$, $p_{3,n_1,n_2}^{\pm}$, $\omega_{n_1,n_2}^{\pm}$ are fixed scalars. Indeed, the case of a general load having the stipulated symmetry can be covered by superposition (e.g., Timoshenko and Woinowsky-Krieger, 1982). Hence, in what follows it is assumed that $p_1^{\pm}$, $p_2^{\pm}$, $p_3^{\pm}$, and $\omega^{\pm}$ have the form specified in Eqs. (19).

6 A Separate-Variables Three-Dimensional Solution

On account of the present rectangular geometry, it is natural to search for a three-dimensional solution of the plate problem such that each of the four unknown functions s_1, s_2, s_3 and ϕ is a product of functions, each depending on one variable only. Indeed, let the displacement and the potential field have the expressions

$$\begin{cases} s_1(x_1, x_2, x_3) = S_1(x_3) \sin q_1 x_1 \cos q_2 x_2 \\ s_2(x_1, x_2, x_3) = S_2(x_3) \cos q_1 x_1 \sin q_2 x_2 \\ s_3(x_1, x_2, x_3) = S_3(x_3) \cos q_1 x_1 \cos q_2 x_2 \\ \phi(x_1, x_2, x_3) = \Phi(x_3) \cos q_1 x_1 \cos q_2 x_2, \end{cases} \tag{20}$$

where

$$q_1 = (2n_1+1)\pi/L_1 \quad \text{and} \quad q_2 = (2n_2+1)\pi/L_2, \tag{21}$$

and S_1, S_2, S_3, Φ are functions to be determined, depending on x_3 only.

As a consequence of this choice, the boundary conditions (17) are trivially satisfied; the field Eqs. (14), using Eqs. (8), become (Tiersten, 1969)

$$\begin{cases} (-C_{11}^{\mathrm{e}} q_1^2 - C_{66}^{\mathrm{e}} q_2^2 + C_{44}^{\mathrm{e}} d^2) S_1 - (C_{12}^{\mathrm{e}} + C_{66}^{\mathrm{e}}) q_1 q_2 S_2 \\ \qquad - (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}}) q_1 dS_3 - (L_{31} + L_{15}) q_1 d\Phi = 0 \\ -(C_{12}^{\mathrm{e}} + C_{66}^{\mathrm{e}}) q_1 q_2 S_1 + (-C_{11}^{\mathrm{e}} q_2^2 - C_{66}^{\mathrm{e}} q_1^2 + C_{44}^{\mathrm{e}} d^2) S_2 \\ \qquad - (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}}) q_2 dS_3 - (L_{31} + L_{15}) q_2 d\Phi = 0 \\ (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}})(q_1 dS_1 + q_2 dS_2) + [-C_{44}^{\mathrm{e}}(q_1^2 + q_2^2) + C_{33}^{\mathrm{e}} d^2] S_3 \\ \qquad + [-L_{15}(q_1^2 + q_2^2) + L_{33} d^2]\Phi = 0 \\ (L_{31} + L_{15})(q_1 dS_1 + q_2 dS_2) + [-L_{15}(q_1^2 + q_2^2) + L_{33} d^2] S_3 \\ \qquad + [\epsilon_{11}^{\mathrm{E}}(q_1^2 + q_2^2) - \epsilon_{33}^{\mathrm{E}} d^2]\Phi = 0 \end{cases} \tag{22}$$

where d denotes the differential operator with respect to x_3; the boundary conditions (18), taking into account Eqs. (19), become

$$\begin{cases} C_{44}^{\mathrm{e}}[dS_1(\pm H/2) - q_1 S_3(\pm H/2)] \\ \qquad - L_{15} q_1 \Phi(\pm H/2) = \pm p_{1,n_1,n_2}^{\pm} \\ C_{44}^{\mathrm{e}}[dS_2(\pm H/2) - q_2 S_3(\pm H/2)] \\ \qquad - L_{15} q_2 \Phi(\pm H/2) = \pm p_{2,n_1,n_2}^{\pm} \\ C_{13}^{\mathrm{e}}[q_1 S_1(\pm H/2) + q_2 S_2(\pm H/2)] + C_{33}^{\mathrm{e}} dS_3(\pm H/2) \\ \qquad + L_{33} d\Phi(\pm H/2) = \pm p_{3,n_1,n_2}^{\pm} \\ L_{31}[q_1 S_1(\pm H/2) + q_2 S_2(\pm H/2)] + L_{33} dS_3(\pm H/2) \\ \qquad - \epsilon_{33}^{\mathrm{E}} d\Phi(\pm H/2) = \mp \omega_{n_1,n_2}^{\pm}. \end{cases} \tag{23}$$

Hence, the plate problem is reduced to the boundary value problem (22)–(23), governed by an *ordinary* differential linear system in the unknown functions S_1, S_2, S_3, Φ.

For the sake of simplifying the problem, the following change of variables is made:

$$U := \nu_1 S_1 + \nu_2 S_2, \; V := -\nu_2 S_1 + \nu_1 S_2, \; X_3 := x_3/l, \tag{24}$$

where

$$l = 1/\sqrt{q_1^2 + q_2^2}, \quad \nu_1 = q_1 l \quad \text{and} \quad \nu_2 = q_2 l. \tag{25}$$

The quantity l, depending upon L_1, L_2, n_1, n_2 has the meaning of a characteristic in-plane dimension of the plate. It is a simple matter to verify that the system (22) is decoupled into the system

$$\begin{cases} (-C_{11}^{\mathrm{e}} + C_{44}^{\mathrm{e}} D^2) U - (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}}) D S_3 \\ \qquad - (L_{31} + L_{15}) D\Phi = 0 \\ (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}}) DU + (-C_{44}^{\mathrm{e}} + C_{33}^{\mathrm{e}} D^2) S_3 \\ \qquad + (-L_{15} + L_{33} D^2)\Phi = 0 \\ (L_{31} + L_{15}) DU + (-L_{15} + L_{33} D^2) S_3 \\ \qquad + (\epsilon_{11}^{\mathrm{E}} - \epsilon_{33}^{\mathrm{E}} D^2)\Phi = 0, \end{cases} \tag{26}$$

and the equation

$$(-C_{66}^{\mathrm{e}} + C_{44}^{\mathrm{e}} D^2) V = 0, \tag{27}$$

where $D = ld$ denotes the differential operator with respect to X_3.

It is worth noting that the coefficients of Eqs. (26) and (27) depend only upon the material constants and are independent of q_1 and q_2: hence, the general solution of Eqs. (26) and (27) can be found once for a given material, and then used for a rectangular plate made of that material, independently of the dimensions or the load.

The characteristic equation of the system (26), in the unknown λ, is

$$\det \begin{bmatrix} -C_{11}^{\mathrm{e}} + C_{44}^{\mathrm{e}}\lambda^2 & -(C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}})\lambda & -(L_{31} + L_{15})\lambda \\ (C_{13}^{\mathrm{e}} + C_{44}^{\mathrm{e}})\lambda & -C_{44}^{\mathrm{e}} + C_{33}^{\mathrm{e}}\lambda^2 & -L_{15} + L_{33}\lambda^2 \\ (L_{31} + L_{15})\lambda & -L_{15} + L_{33}\lambda^2 & \epsilon_{11}^{\mathrm{E}} - \epsilon_{33}^{\mathrm{E}}\lambda^2 \end{bmatrix} = 0, \tag{28}$$

where "det" denotes the determinant. In a more explicit form, Eq. (28) can be written as

$$-a_6\lambda^6 + a_4\lambda^4 - a_2\lambda^2 + a_0 = 0, \tag{29}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$\begin{cases} a_6 = C^e_{44}(C^e_{33}\epsilon^E_{33} + L^2_{33}), \\ a_4 = C^e_{44}C^e_{33}\epsilon^E_{11} + (C^e_{44})^2\epsilon^E_{33} + C^e_{11}C^e_{33}\epsilon^E_{33} + L^2_{33}C^e_{11} \\ \quad + 2L_{33}L_{15}C^e_{44} - 2(C^e_{13} + C^e_{44})(L_{31} + L_{15})L_{33} \\ \quad + (L_{31} + L_{15})^2C^e_{33} - (C^e_{13} + C^e_{44})^2\epsilon^E_{33}, \\ a_2 = C^e_{33}C^e_{11}\epsilon^E_{11} + (C^e_{44})^2\epsilon^E_{11} + C^e_{11}C^e_{44}\epsilon^E_{33} + L^2_{15}C^e_{44} \\ \quad + 2L_{33}L_{15}C^e_{11} - 2(C^e_{13} + C^e_{44})(L_{31} + L_{15})L_{15} \\ \quad + (L_{31} + L_{15})^2C^e_{44} - (C^e_{13} + C^e_{44})^2\epsilon^E_{11}, \\ a_0 = C^e_{11}(C^e_{44}\epsilon^E_{11} + L^2_{15}). \end{cases} \tag{30}$$

For the sake of simplicity, it is assumed that Eq. (29) has six distinct roots:

$$\lambda_1, -\lambda_1, \lambda_2, -\lambda_2, \lambda_3, -\lambda_3.$$

If $[U(X_3), S_3(X_3), \Phi(X_3)]$ is a solution of the system (26), then $[-U(-X_3), S_3(-X_3), \Phi(-X_3)]$ is a solution, too. Hence, the general solution of the system (26) can be given the form

$$\begin{cases} U(X_3) = U^e(X_3) + U^o(X_3) \\ S_3(X_3) = S^o_3(X_3) + S^e_3(X_3) \\ \Phi(X_3) = \Phi^o(X_3) + \Phi^e(X_3), \end{cases} \tag{31}$$

where

$$\begin{cases} U^e(X_3) = \sum_{i=1}^{3} A_i\alpha_i \cosh(\lambda_i X_3) \\ S^o_3(X_3) = \sum_{i=1}^{3} A_i\beta_i \sinh(\lambda_i X_3) \\ \Phi^o(X_3) = \sum_{i=1}^{3} A_i\gamma_i \sinh(\lambda_i X_3) \end{cases}$$

and

$$\begin{cases} U^o(X_3) = \sum_{i=1}^{3} B_i\alpha_i \sinh(\lambda_i X_3) \\ S^e_3(X_3) = \sum_{i=1}^{3} B_i\beta_i \cosh(\lambda_i X_3) \\ \Phi^e(X_3) = \sum_{i=1}^{3} B_i\gamma_i \cosh(\lambda_i X_3). \end{cases} \tag{32}$$

The superscripts "e" and "o" denote even and odd functions of X_3, respectively; A_i, B_i are arbitrary dimensionless (complex) constants. The vector $(\alpha_i, \beta_i, \gamma_i)$ spans the null space of the matrix at the left-hand side of Eq. (28) when λ is given the value λ_i, for $i = 1 \ldots 3$ (of course, this vector is defined up to a scalar multiple). In other words, the constants α_i, β_i, γ_i, satisfy the following equations, where $i = 1 \ldots 3$ is unsummed:

$$\begin{cases} (-C^e_{11} + C^e_{44}\lambda_i^2)\alpha_i - (C^e_{13} + C^e_{44})\lambda_i\beta_i \\ \qquad - (L_{31} + L_{15})\lambda_i\gamma_i = 0 \\ (C^e_{13} + C^e_{44})\lambda_i\alpha_i + (-C^e_{44} + C^e_{33}\lambda_i^2)\beta_i \\ \qquad + (-L_{15} + L_{33}\lambda_i^2)\gamma_i = 0 \\ (L_{31} + L_{15})\lambda_i\alpha_i + (-L_{15} + L_{33}\lambda_i^2)\beta: + (\epsilon^E_{11} - \epsilon^E_{33}\lambda_i^2)\gamma_i = 0. \end{cases} \tag{33}$$

Equation (27) has the general solution

$$V(X_3) = V^e(X_3) + V^o(X_3), \tag{34}$$

where

$$V^e(X_3) = A_4 1 \cosh(\lambda_4 X_3) \quad \text{and}$$
$$V^o(X_3) = B_4 1 \sinh(\lambda_4 X_3), \tag{35}$$

$\lambda_4 = \sqrt{C^e_{66}/C^e_{44}}$, A_4, B_4 are arbitrary dimensionless constants, and 1 has the physical dimension of a length.

For the piezoelectric materials considered in Table 1, the values of the constants λ_i, α_i, β_i, and γ_i are given in Table 2, in the units of the International System (j denotes the imaginary unit).

The boundary conditions (23), according to the change of variables (24), are decoupled into

(i) a system of three equations involving U^e, S^o_3, Φ^o:

$$\begin{cases} C^e_{44}DU^e(h/2) - C^e_{44}S^o_3(h/2) - L_{15}\Phi^o(h/2) = lp^e \\ C^e_{13}U^e(h/2) + C^e_{33}DS^o_3(h/2) + L_{33}D\Phi^o(h/2) = lp^o_3 \\ L_{31}U^e(h/2) + L_{33}DS^o_3(h/2) - \epsilon^E_{33}D\Phi^o(h/2) = -l\omega^o; \end{cases} \tag{36}$$

(ii) a system of three equations involving U^o, S^e_3, Φ^e:

$$\begin{cases} C^e_{44}DU^o(h/2) - C^e_{44}S^e_3(h/2) - L_{15}\Phi^e(h/2) = lp^o \\ C^e_{13}U^o(h/2) + C^e_{33}DS^e_3(h/2) + L_{33}D\Phi^e(h/2) = lp^e_3 \\ L_{31}U^o(h/2) + L_{33}DS^e_3(h/2) - \epsilon^E_{33}D\Phi^e(h/2) = -l\omega^e; \end{cases} \tag{37}$$

(iii) one equation involving V^e:

$$C^e_{44}DV^e(h/2) = lq^e; \tag{38}$$

(iv) one equation involving V^o:

$$C^e_{44}DV^o(h/2) = lq^o, \tag{39}$$

where the plate thickness aspect ratio (i.e., the ratio between the thickness H and the characteristic in-plane dimension l of the plate) is introduced:

$$h := H/l \tag{40}$$

and the following positions are made:

$$\begin{aligned} p^e_1 &= \tfrac{1}{2}(p^+_{1,n_1,n_2} + p^-_{1,n_1,n_2}) & p^o_1 &= \tfrac{1}{2}(p^+_{1,n_1,n_2} - p^-_{1,n_1,n_2}) \\ p^e_2 &= \tfrac{1}{2}(p^+_{2,n_1,n_2} + p^-_{2,n_1,n_2}) & p^o_2 &= \tfrac{1}{2}(p^+_{2,n_1,n_2} - p^-_{2,n_1,n_2}) \\ p^e_3 &= \tfrac{1}{2}(p^+_{3,n_1,n_2} + p^-_{3,n_1,n_2}) & p^o_3 &= \tfrac{1}{2}(p^+_{3,n_1,n_2} - p^-_{3,n_1,n_2}) \\ \omega^e &= \tfrac{1}{2}(\omega^+_{n_1,n_2} + \omega^-_{n_1,n_2}) & \omega^o &= \tfrac{1}{2}(\omega^+_{n_1,n_2} - \omega^-_{n_1,n_2}) \\ p^e &= +\nu_1 p^e_1 + \nu_2 p^e_2 & p^o &= +\nu_1 p^o_1 + \nu_2 p^o_2 \\ q^e &= -\nu_2 p^e_1 + \nu_1 p^e_2 & q^o &= -\nu_2 p^o_1 + \nu_1 p^o_2. \end{aligned} \tag{41}$$

In a more explicit form, Eqs. (36) and (37) can be respectively rewritten as

$$\begin{cases} \sum_{i=1}^{3} A_i t_i \sinh(\lambda_i h/2) = lp^e \\ \sum_{i=1}^{3} A_i r_i \cosh(\lambda_i h/2) = lp^o_3 \\ \sum_{i=1}^{3} A_i z_i \cosh(\lambda_i h/2) = -l\omega^o \end{cases}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 2 The material constants λ_i, α_i, β_i, and γ_i

	CdS	PZT-4	PZT-5H	PZT-8	C-24
λ_1	$1.665 + j0.$	$1.069 + j0.200$	$1.029 + j0.415$	$1.044 + j0.242$	$1.032 + j0.127$
α_1 [m]	$-0.913 + j0.$	$0.254 - j0.418$	$0.632 - j0.248$	$0.594 - j0.183$	$0.441 - j0.516$
β_1 [m]	$0.409 + j0.$	$-0.080 + j0.459$	$-0.329 + j0.464$	$-0.439 + j0.350$	$-0.300 + j0.630$
γ_1 [V]	$0.008 + j0.$	$-0.659 + j0.330$	$-0.327 + j0.330$	$-0.407 + j0.364$	$0.114 + j0.199$
λ_2	$0.591 + j0.$	$1.069 - j0.200$	$1.029 - j0.415$	$1.044 - j0.242$	$1.032 - j0.127$
α_2 [m]	$0.415 + j0.$	$0.254 + j0.418$	$0.632 + j0.248$	$0.594 + j0.183$	$0.441 + j0.516$
β_2 [m]	$-0.910 + j0.$	$-0.080 - j0.459$	$-0.329 - j0.464$	$-0.439 - j0.350$	$-0.300 - j0.630$
γ_2 [V]	$0.009 + j0.$	$-0.659 - j0.330$	$-0.327 - j0.330$	$-0.407 - j0.364$	$0.114 - j0.199$
λ_3	$0.973 + j0.$	$1.204 + j0.$	$1.053 + j0.$	$1.195 + j0.$	$1.065 + j0.$
α_3 [m]	$0.047 + j0.$	$-0.696 + j0.$	$-0.090 + j0.$	$-0.018 + j0.$	$-0.024 + j0.$
β_3 [m]	$-0.050 + j0.$	$0.617 + j0.$	$0.176 + j0.$	$-0.048 + j0.$	$-0.040 + j0.$
γ_3 [V]	$0.998 + j0.$	$-0.366 + j0.$	$-0.980 + j0.$	$0.999 + j0.$	$0.999 + j0.$
λ_4	$1.042 + j0.$	$1.093 + j0.$	$1.005 + j0.$	$1.035 + j0.$	$1.013 + j0.$

$$\text{and} \quad \begin{cases} \sum_{i=1}^{3} B_i t_i \cosh(\lambda_i h/2) = l p^o, \\ \sum_{i=1}^{3} B_i r_i \sinh(\lambda_i h/2) = l p_3^e \\ \sum_{i=1}^{3} B_i z_i \sinh(\lambda_i h/2) = -l \omega^e, \end{cases} \tag{42}$$

where

$$\begin{cases} t_i = C_{44}^{\mathrm{e}} \alpha_i \lambda_i - C_{44}^{\mathrm{e}} \beta_i - L_{15} \gamma_i \\ r_i = C_{13}^{\mathrm{e}} \alpha_i + C_{33}^{\mathrm{e}} \beta_i \lambda_i + L_{33} \gamma_i \lambda_i \\ z_i = L_{31} \alpha_i + L_{33} \beta_i \lambda_i - \epsilon_{33}^{\mathrm{E}} \gamma_i \lambda_i \end{cases} \quad (i = 1 \ldots 3, \text{unsummed}). \tag{43}$$

Hence, they allow to determine the constants A_1, A_2, A_3 and B_1, B_2, B_3, respectively. Analogously, Eqs. (38) and (39) can be, respectively, rewritten as

$$C_{44}^{\mathrm{e}} A_4 \lambda_4 \sinh(\lambda_4 h/2) = l q^e \quad \text{and}$$

$$C_{44}^{\mathrm{e}} B_4 \lambda_4 \cosh(\lambda_4 h/2) = l q^o, \tag{44}$$

and allow to determine the constants A_4 and B_4, respectively.

The functions V, U, S_3, Φ are determined by Eqs. (31) and (34); these functions turn out to be real, in spite of the fact that, in general, A_i, B_i are complex constants. Then, the displacement field is obtained by Eqs. (20) and (24):

$$\begin{cases} s_1 = (\nu_1 U - \nu_2 V) \sin q_1 x_1 \cos q_2 x_2 \\ s_2 = (\nu_2 U + \nu_1 V) \cos q_1 x_1 \sin q_2 x_2 \\ s_3 = S_3 \cos q_1 x_1 \cos q_2 x_2 \\ \phi = \Phi \cos q_1 x_1 \cos q_2 x_2; \end{cases} \tag{45}$$

the strain field is obtained by Eq. (10):

$$\begin{cases} E_1 = \nu_1 / l (\nu_1 U - \nu_2 V) \cos q_1 x_1 \cos q_2 x_2 \\ E_2 = \nu_2 / l (\nu_2 U + \nu_1 V) \cos q_1 x_1 \cos q_2 x_2 \\ E_3 = 1/l D S_3 \cos q_1 x_1 \cos q_2 x_2 \\ 2E_4 = 1/l [\nu_2 (DU - S_3) + \nu_1 DV] \cos q_1 x_1 \sin q_2 x_2 \\ 2E_5 = 1/l [\nu_1 (DU - S_3) - \nu_2 DV] \sin q_1 x_1 \cos q_2 x_2 \\ 2E_6 = 1/l [-2\nu_1 \nu_2 U + (\nu_2^2 - \nu_1^2) V] \sin q_1 x_1 \sin q_2 x_2; \end{cases} \tag{46}$$

the electric field is obtained by Eq. (11):

$$\begin{cases} e_1 = \nu_1 / l \Phi \sin q_1 x_1 \cos q_2 x_2 \\ e_2 = \nu_2 / l \Phi \cos q_1 x_1 \sin q_2 x_2 \\ e_3 = -1/l D \Phi \cos q_1 x_1 \cos q_2 x_2; \end{cases} \tag{47}$$

the stress field is obtained by Eqs. (5) and (8):

$$\begin{cases} T_1 = 1/l [Q - 2\nu_2 C_{66}^{\mathrm{e}} (\nu_2 U + \nu_1 V)] \cos q_1 x_1 \cos q_2 x_2 \\ T_2 = 1/l [Q - 2\nu_1 C_{66}^{\mathrm{e}} (\nu_1 U - \nu_2 V)] \cos q_1 x_1 \cos q_2 x_2 \\ T_3 = R/l \cos q_1 x_1 \cos q_2 x_2 \\ T_4 = 1/l (\nu_2 W + \nu_1 C_{44}^{\mathrm{e}} DV) \cos q_1 x_1 \sin q_2 x_2 \\ T_5 = 1/l (\nu_1 W - \nu_2 C_{44}^{\mathrm{e}} DV) \sin q_1 x_1 \cos q_2 x_2 \\ T_6 = C_{66}^{\mathrm{e}} / l [-2\nu_1 \nu_2 U + (\nu_2^2 - \nu_1^2) V] \sin q_1 x_1 \sin q_2 x_2; \end{cases} \tag{48}$$

and the electric displacement field is obtained by Eqs. (5) and (8), too,

$$\begin{cases} d_1 = 1/l (\nu_1 Y - L_{15} \nu_2 DV) \sin q_1 x_1 \cos q_2 x_2 \\ d_2 = 1/l (\nu_2 Y + L_{15} \nu_1 DV) \cos q_1 x_1 \sin q_2 x_2 \\ d_3 = Z/l \cos q_1 x_1 \cos q_2 x_2, \end{cases} \tag{49}$$

where the following positions have been made:

$$\begin{cases} Q = C_{11}^{\mathrm{e}} U + C_{13}^{\mathrm{e}} D S_3 + L_{31} D \Phi \\ R = C_{13}^{\mathrm{e}} U + C_{33}^{\mathrm{e}} D S_3 + L_{33} D \Phi \\ W = C_{44}^{\mathrm{e}} DU - C_{44}^{\mathrm{e}} S_3 - L_{15} \Phi \\ Y = L_{15} DU - L_{15} S_3 + \epsilon_{11}^{\mathrm{E}} \Phi \\ Z = L_{31} U + L_{33} D S_3 - \epsilon_{33}^{\mathrm{E}} D \Phi. \end{cases} \tag{50}$$

Needless to say that if

$$\omega_{n_1,n_2}^{+} = \omega_{n_1,n_2}^{-}, \quad p_{3,n_1,n_2}^{+} = p_{3,n_1,n_2}^{-},$$
$$p_{1,n_2,n_2}^{+} = -p_{1,n_1,n_2}^{-}, \quad p_{2,n_1,n_2}^{+} = -p_{2,n_1,n_2}^{-}, \tag{51}$$

then $\omega^o = 0$, $p_3^o = 0$, $p^e = 0$, $q^e = 0$: hence the constants A_i turn out to be zero and $V^e = 0$, $U^e = 0$, $S_3^o = 0$, $\Phi^o = 0$. As a consequence, s_3, ϕ, E_4, E_5, e_1, e_2, T_4, T_5, d_1 and d_2 are even functions of the x_3 variable, while s_1, s_2, E_1, E_2, E_3, E_6, e_3, T_1, T_2, T_3, T_6 and d_3 are odd functions of the x_3 variable: this amounts to say that the plate has a bending behavior. Conversely, if

$$\omega_{n_1,n_2}^{+} = -\omega_{n_1,n_2}^{-}, \quad p_{3,n_1,n_2}^{+} = -p_{3,n_1,n_2}^{-},$$
$$p_{1,n_1,n_2}^{+} = p_{1,n_1,n_2}^{-}, \quad p_{2,n_1,n_2}^{+} = p_{2,n_1,n_2}^{-}, \tag{52}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 3 Auxiliary material constants

	CdS	PZT-4	PZT-5H	PZT-8	C-24
C [GPa]	63.1	120.	96.2	117.	143.
G [GPa]	59.7	153.	102.	163.	161.
K [GPa]	1.80	−7.66	9.49	4.33	0.236
F [C/m^2]	0.458	25.1	36.3	20.4	0.622
M [C/m^2]	0.529	36.3	46.5	29.9	10.9
N [C/m^2]	−1.34	34.6	34.8	18.4	9.43
ϵ [nF/m]	79.8	12.8	27.6	11.4	2.09
μ	0.543	0.388	0.456	0.378	0.222
σ	0.967	0.924	0.814	0.870	0.821
ρ	1.00	0.793	0.767	0.841	0.992

then $\omega^e = 0$, $p_3^e = 0$, $p^o = 0$, $q^o = 0$. Hence, the constants B_i turn out to be zero and the previously stated evenness and oddness properties exchange: in this case, the plate exhibits a membrane behavior.

7 The Limit of the Three-Dimensional Solution as the Plate Thickness Aspect Ratio Approaches Zero

In this section the limit of the three-dimensional solution (45) – (49) as the plate thickness aspect ratio h approaches zero is explicitly carried out, by performing the expansion into a power series of h of the involved mechanical and electric quantities, and retaining only the lowest-order terms. The result obtained is denoted as *thin-plate limit* of the three-dimensional solution.

In order to preserve, in the limit, the dependence upon the variable in the thickness direction, the rescaled variable

$$\zeta := \frac{X_3}{h} = \frac{x_3}{H} \tag{53}$$

is introduced ($\zeta = 0$ on the middle plane of the plate, and $\zeta = \pm\frac{1}{2}$ on the upper or lower face).

Needless to say that, in spite of the fact that the mechanical and electric quantities are related by Eqs. (10), (11), and (12), the lowest-order terms of their expansions into a power series of h needn't be related by those equations.

The calculations are based on the expansion into a power series of h of the solutions of Eqs. (42) and (44). The constants λ_i, α_i, β_i, γ_i are irrational functions of the material constants introduced in Eqs. (8). However, by taking into account the fact that, for $i = 1 \ldots 3$, λ_i solves Eq. (29) and α_i, β_i, γ_i satisfy Eqs. (33), it is possible to prevent the constants λ_i, α_i, β_i, γ_i by appearing explicitly in the thin-plate limit. Some auxiliary material constants are used instead, depending rationally on the ones introduced in Eq. (8):

$$C := C_{11}^e - \frac{\epsilon_{33}^E (C_{13}^e)^2 + 2L_{31}L_{33}C_{13}^e - L_{31}^2 C_{33}^e}{C_{33}^e \epsilon_{33}^E + L_{33}^2}$$

$$\epsilon := \epsilon_{11}^E + \frac{L_{15}^2}{C_{44}^e}$$

$$G := \epsilon \frac{C_{11}^e C_{33}^e - (C_{13}^e)^2}{C_{33}^e \epsilon_{33}^E + L_{33}^2} \quad F := \epsilon \frac{L_{33}C_{13}^e - C_{33}^e L_{31}}{C_{33}^e \epsilon_{33}^E + L_{33}^2}$$

$$M := \epsilon \frac{L_{33}C_{11}^e - C_{13}^e L_{31}}{C_{33}^e \epsilon_{33}^E + L_{33}^2} \quad N := \frac{L_{15}C - C_{44}^e F}{C_{44}^e}$$

$$K := \frac{MC - NG}{M} \quad \mu := \frac{C_{13}^e \epsilon_{33}^E + L_{33}L_{31}}{C_{33}^e \epsilon_{33}^E + L_{33}^2}$$

$$\sigma := \frac{C_{11}^e \epsilon_{33}^E + L_{31}^2}{C_{33}^e \epsilon_{33}^E + L_{33}^2} \quad \rho := 1 - \frac{L_{15}F}{\epsilon C}. \tag{54}$$

For the piezoelectric materials considered in Table 1, the values of these constants are given in Table 3, in the units of the International System.

Only one at a time of the eight quantities p_3^e, p^o, ω^e, q^o, p_3^o, p^e, ω^o, q^e is assumed to be different from zero: hence, eight different cases are considered. Any load condition can be covered by superposition of these eight basic cases. The results obtained if either one of the quantities p_3^e, p^o, ω^e, q^o is different from zero are reported in Table 4 (in this case, the plate has a bending behavior). The results obtained if either one of the quantities p_3^o, p^e, ω^o, q^e is different from zero are reported in Table 5 (in this case, the plate has a membrane behavior).

For a square plate of edge $L_1 = L_2 = A$ and thickness H, some useful formulas can be easily carried out by the results given in Tables 4 and 5. The limits, as the plate thickness aspect ratio approaches zero, of some mechanical and electric quantities of primary interest in technical applications are found. These results should be recovered by applying any consistent theory of thin piezoelectric plates.

Let $f(x_1, x_2) := \cos(\pi x_1/A) \cos(\pi x_2/A)$. In Table 6 there are reported the deflection w_m, the electric potential Φm, and the bending moment M_{1m} at the center of the plate:

$$w_m := s_3(A/2, A/2, 0)$$

$$\Phi_m := \phi(A/2, A/2, H/2)$$

$$M_{1m} := \int_{-H/2}^{H/2} x_3 T_1(A/2, A/2, x_3)\, dx_3 \tag{55}$$

when the plate is subjected only to a normal sinusoidal load of largest intensity p_0 ($p_3^{\pm} = p_0 f(x_1, x_2)/2$), or to a normal uniform load of intensity p_0 ($p_3^{\pm} = p_0/2$), or to a sinusoidal surface charge of largest density ω_0 ($\omega^{\pm} = \omega_0 f(x_1, x_2)/2$), or to a uniform surface charge of density ω_0 ($\omega^{\pm} = \omega_0/2$). In these cases the plate has a bending behavior. The following position has been made:

$$\xi = C_{66}^e/C. \tag{56}$$

It is pointed out that the results obtained for w_m when a sinusoidal or uniform load p_0 is applied, formally coincide with the results about an elastic (i.e., nonpiezoelectric) plate given by Lekhnitskii (1968), if the quantity $H^3C/12$ is replaced by $H^3/12[C_{11}^e - (C_{13}^e)^2/C_{33}^e]$. The latter quantity is the bending stiffness of an elastic plate. Noting that, if no piezoelectric coupling exists, the expression for C given in Eqs. (54) reduces just to $C_{11}^e - (C_{13}^e)^2/C_{33}^e$, it is natural to call the quantity $H^3C/12$ "bending stiffness" of a piezoelectric plate.

In Table 7 there are reported the variation of the thickness ΔH_m, the difference of electric potential $\Delta\Phi_m$, and the membrane force N_{1m} at the center of the plate:

$$\Delta H_m := s_3(A/2, A/2, H/2) - s_3(A/2, A/2, -H/2)$$

$$\Delta\Phi_m := \phi(A/2, A/2, H/2) - \phi(A/2, A/2, -H/2)$$

$$N_{1m} := \int_{-H/2}^{H/2} T_1(A/2, A/2, x_3)\, dx_3 \tag{57}$$

when the plate is subjected on the upper and lower face only to a normal sinusoidal load of largest intensity p_0 ($p_3^{\pm} = \pm p_0 f(x_1, x_2)$), or to a normal uniform load of intensity p_0 ($p_3^{\pm} = \pm p_0$), or to a sinusoidal surface charge of largest density ω_0 ($\omega^{\pm} = \pm\omega_0 f(x_1, x_2)$), or to a uniform surface charge of density ω_0 ($\omega^{\pm} = \pm\omega_0$). In these cases the plate has a membrane behavior. It is emphasized that both the sinusoidal and the uniform distribution produce the same results.

The results provided by the zeroth-order theory of Maugin and Attou (1990), based on the scaling theory of Ciarlet and Destuynder (1979) do not agree with the thin-plate limit obtained here. Indeed, the deflection of a plate according to the zeroth order theory of Maugin and Attou is governed by an equation (Eq. (7.10) in Maugin and Attou, 1990) which doesn't

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 4 Thin-plate limit of the three-dimensional solution. Bending behavior.

	p_3^e	p^o	ω^e	q^o
U	$\frac{24\ell p_3^e}{h^2 C}\zeta$	$\frac{12\ell p^o}{hC}\zeta$	$2\ell\omega^e\frac{F}{\epsilon C}\zeta$	0
S_3	$\frac{24\ell p_3^e}{h^3 C}$	$\frac{12\ell p^o}{h^2 C}$	$\frac{-2\ell\omega^e}{h}\frac{N}{\epsilon C}$	0
Φ	$\frac{-2\ell p_3^e}{h}\frac{F}{\epsilon C}[-\frac{N}{F}-\frac{3}{2}+6\zeta^2]$	$-\ell p^o\frac{F}{\epsilon C}(6\zeta^2-\frac{1}{2})$	$\frac{2\ell\omega^e}{h\epsilon}$	0
DS_3	$\frac{-24\ell p_3^e}{h^2}\frac{\mu}{C}\zeta$	$\frac{-12\ell p^o}{h}\frac{\mu}{C}\zeta$	$-2\ell\omega^e\frac{M}{\epsilon C}\zeta$	0
$D\Phi$	$\frac{-24\ell p_3^e}{h^2}\frac{F}{\epsilon C}\zeta$	$\frac{-12\ell p^o}{h}\frac{F}{\epsilon C}\zeta$	$\cdot 2\ell\omega^e\frac{G}{\epsilon C}\zeta$	0
$DU-S_3$	$\frac{2\ell p_3^e}{hC_{44}^e}[\frac{-\epsilon_{11}}{\epsilon}+\rho(6\zeta^2-\frac{1}{2})]$	$\frac{\ell p^o\rho}{C_{44}^e}(6\zeta^2-\frac{1}{2})$	$\frac{2\ell\omega^e}{h}\frac{L_{15}}{\epsilon C_{44}^e}$	0
Q	$\frac{24\ell p_3^e}{h^2}\zeta$	$\frac{12\ell p^o}{h}\zeta$	$\frac{-h^2\ell\omega^e}{60}\frac{MK}{\epsilon C}(-3\zeta+20\zeta^3)$	0
R	$\ell p_3^e(3\zeta-4\zeta^3)$	$\frac{h\ell p^o}{2}(\zeta-4\zeta^3)$	$\frac{h^4\ell\omega^e}{960}\frac{MK}{\epsilon C}(\zeta-8\zeta^3+16\zeta^5)$	0
W	$\frac{-3\ell p_3^e}{h}(1-4\zeta^2)$	$\ell p^o(6\zeta^2-\frac{1}{2})$	$\frac{-h^3\ell\omega^e}{960}\frac{MK}{\epsilon C}(1-24\zeta^2+80\zeta^4)$	0
Y	$\frac{2\ell p_3^e}{h}\frac{N}{C}(6\zeta^2-\frac{1}{2})$	$\ell p^o\frac{N}{C}(6\zeta^2-\frac{1}{2})$	$\frac{2\ell\omega^e}{h}$	0
Z	$\ell p_3^e\frac{N}{C}(\zeta-4\zeta^3)$	$\frac{h\ell p^o}{2}\frac{N}{C}(\zeta-4\zeta^3)$	$-2\ell\omega^e\zeta$	0
V	0	0	0	$\frac{h\ell q^o}{C_{44}^e}\zeta$
DV	0	0	0	$\frac{\ell q^o}{C_{44}^e}$

Table 5 Thin-plate limit of the three-dimensional solution. Membrane behavior.

	p_3^o	p^e	ω^o	q^e
U	$-\ell p_3^o\frac{\mu}{C}$	$\frac{2\ell p^e}{hC}$	$\ell\omega^o\frac{F}{\epsilon C}$	0
S_3	$\ell h p_3^o\frac{\sigma}{C}\zeta$	$-2\ell p^e\frac{\mu}{C}\zeta$	$-\ell h\omega^o\frac{M}{\epsilon C}\zeta$	0
Φ	$\ell h p_3^o\frac{M}{\epsilon C}\zeta$	$-2\ell p^e\frac{F}{\epsilon C}\zeta$	$\ell h\omega^o\frac{G}{\epsilon C}\zeta$	0
DS_3	$\ell p_3^o\frac{\sigma}{C}$	$\frac{-2\ell p^e}{h}\frac{\mu}{C}$	$-\ell\omega^o\frac{M}{\epsilon C}$	0
$D\Phi$	$\ell p_3^o\frac{M}{\epsilon C}$	$\frac{-2\ell p^e}{h}\frac{F}{\epsilon C}$	$\ell\omega^o\frac{G}{\epsilon C}$	0
$DU-S_3$	$\frac{\ell h p_3^o L_{15}}{C_{44}^e}\frac{M}{C\epsilon}\zeta$	$\frac{2\ell p^e\rho}{C_{44}^e}\zeta$	$\frac{\ell h\omega^o L_{15}}{C_{44}^e}\frac{G}{\epsilon C}\zeta$	0
Q	$\frac{p_3^o h^2\ell}{12}[\sigma+\frac{NM}{\epsilon C}](6\zeta^2-\frac{1}{2})$	$\frac{2\ell p^e}{h}$	$\frac{-h^2\ell\omega^o}{12}\frac{MK}{\epsilon C}(6\zeta^2-\frac{1}{2})$	0
R	ℓp_3^o	$\frac{h\ell p^e}{4}(1-4\zeta^2)$	$\frac{h^4\ell\omega^o}{384}\frac{MK}{\epsilon C}(1-8\zeta^2+16\zeta^4)$	0
W	$\frac{-p_3^o\ell h^3}{24}[\sigma+\frac{NM}{\epsilon C}](\zeta-4\zeta^3)$	$2\ell p^e\zeta$	$\frac{h^3\ell\omega^o}{24}\frac{MK}{\epsilon C}(\zeta-4\zeta^3)$	0
Y	$\ell h p_3^o\frac{M}{C}\zeta$	$2\ell p^e\frac{N}{C}\zeta$	$h\ell\omega^o\frac{G}{C}\zeta$	0
Z	$\frac{h^2 p_3^o\ell}{8}\frac{M}{C}(1-4\zeta^2)$	$\frac{h p^e\ell}{4}\frac{N}{C}(1-4\zeta^2)$	$-\ell\omega^o$	0
V	0	0	0	$\frac{2\ell q^e}{hC_{66}^e}$
DV	0	0	0	$\frac{2\ell q^e}{C_{44}^e}\zeta$

Table 6 Some formulas for a square thin plate. Bending behavior.

	sinusoidal p_0	uniform p_0	sinusoidal ω_0	uniform ω_0
w_m	$\frac{3A^4}{\pi^4 H^3}\frac{p_0}{C}$	$\frac{4.7485A^4}{\pi^4 H^3}\frac{p_0}{C}$	$\frac{-A^2}{2\pi^2 H}\frac{\omega_0 N}{\epsilon C}$	$\frac{-0.7271A^2}{\pi^2 H}\frac{\omega_0 N}{\epsilon C}$
Φ_m	$\frac{A^2}{2\pi^2 H}\frac{p_0 N}{\epsilon C}$	$\frac{0.7271A^2}{\pi^2 H}\frac{p_0 N}{\epsilon C}$	$\frac{A^2}{2\pi^2 H}\frac{\omega_0}{\epsilon}$	$\frac{0.7271A^2}{\pi^2 H}\frac{\omega_0}{\epsilon}$
M_{1m}	$\frac{A^2 p_0(1-\xi)}{2\pi^2}$	$\frac{0.7271A^2 p_0(1-\xi)}{\pi^2}$	$\frac{-H^2}{12}\frac{\omega_0 F\xi}{\epsilon}$	$\frac{-H^2}{12}\frac{\omega_0 F\xi}{\epsilon}$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 7 Some formulas for a square thin plate. Membrane behavior.

	sinusoidal p_0	uniform p_0	sinusoidal ω_0	uniform ω_0
ΔH_m	$H\frac{p_0\sigma}{C}$	$H\frac{p_0\sigma}{C}$	$-H\frac{\omega_0 M}{\epsilon C}$	$-H\frac{\omega_0 M}{\epsilon C}$
$\Delta\Phi_m$	$H\frac{p_0 M}{\epsilon C}$	$H\frac{p_0 M}{\epsilon C}$	$H\frac{\omega_0 G}{\epsilon C}$	$H\frac{\omega_0 G}{\epsilon C}$
N_{1m}	$Hp_0\mu\xi$	$Hp_0\mu\xi$	$-H\frac{\omega_0\xi F}{\epsilon}$	$-H\frac{\omega_0\xi F}{\epsilon}$

take into account any piezoelectric coupling. In particular, the deflection given by this theory depends only upon the elastic constants and not upon the piezoelectric and dielectric constants, contradicting the results given in Tables 4 and 6.

A different theory of thin piezoelectric plates was proposed by Tzou (1993). He assumes, as a blanket hypothesis, that the only nonvanishing component of the electric field is e_3 (cf. Section 3.2.2 in Tzou, 1993). The thin-plate limit of the three-dimensional solution obtained in this paper, reported in Tables 4 and 5, shows that the ratio between the in-plane components e_1, e_2 and the transverse component e_3 of the electric field is of the order h, unless $\omega^e \neq 0$, when this ratio is of the order h^{-1}. Hence, noting that h is a small quantity for a thin plate, the hypothesis of Tzou turns out to be oversimplifying, at least in the latter case.

A first-order shear deformation theory of piezoelectric plates was proposed by Mindlin (1972). It is based upon the assumptions that the in-plane displacement components s_1 and s_2 and the potential ϕ are linearly variable with respect to x_3, while the out-of-plane displacement component s_3 is constant with respect to x_3. But, the assumption concerning ϕ is not completely satisfactory, since from Table 4 it turns out that even in the case of a thin plate the potential ϕ has a quadratic law of variation with respect to x_3, when $p_3^e \neq 0$ or $p^o \neq 0$.

Therefore, the interest for a more satisfactory theory of piezoelectric plates arises. The results reported in Tables 4 and 5 suggest which ones of the mechanical and electric quantities may be neglected in the formulation of a consistent thin-plate theory: hence, they may represent a useful guide to develop such a theory.

8 A Numerical Investigation

A numerical investigation in the case of a square plate of edge $L_1 = L_2 = A$ and thickness H is performed, in order to evaluate the influence of the thickness-to-side ratio on the three-dimensional solution of the plate problem. Customarily, in the literature the load is applied according to a sinusoidal law (Pagano, 1970; Srinivas and Rao, 1970; Ray et al., 1992). Here a uniform distribution is adopted, more frequent in technical applications. Really, the load is distributed according to the function:

$$\sum_{n_1=0}^{N_1}\sum_{n_2=0}^{N_2}\frac{16}{\pi^2}\frac{(-1)^{n_1+n_2}}{(2n_1+1)(2n_2+1)} \times \cos\frac{(2n_1+1)\pi x_1}{A}\cos\frac{(2n_2+1)\pi x_2}{A}, \quad (58)$$

which converges uniformly to the constant function 1 in compact subsets of the open square $]-A/2, A/2[\times]-A/2, A/2[$

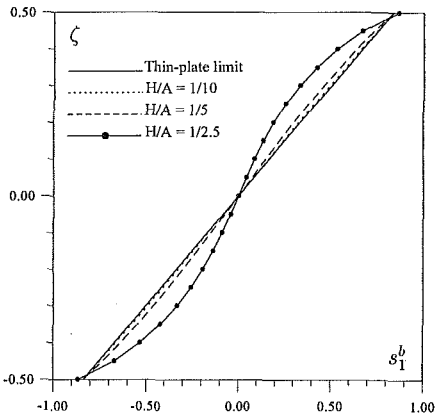


Fig. 1(a) Dimensionless side in-plane displacement s_1^b

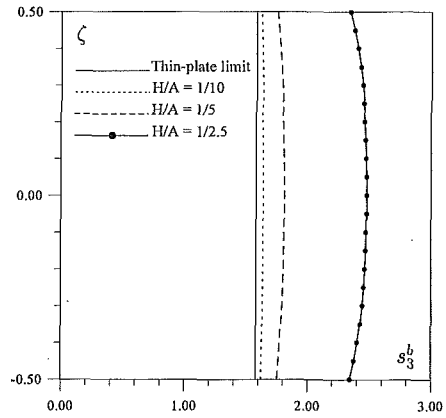


Fig. 1(b) Dimensionless central deflection s_3^b

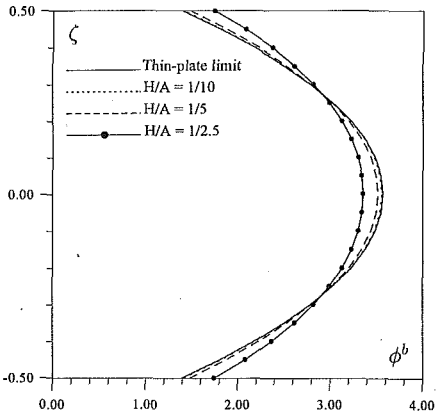


Fig. 1(c) Dimensionless central potential ϕ^b

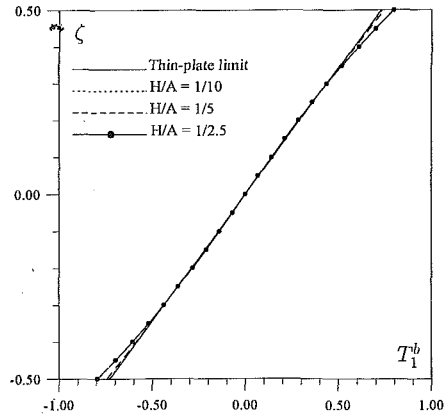


Fig. 1(d) Dimensionless central in-plane stress T_1^b

Fig. 1 Displacements, potential, and stress induced by q_0

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

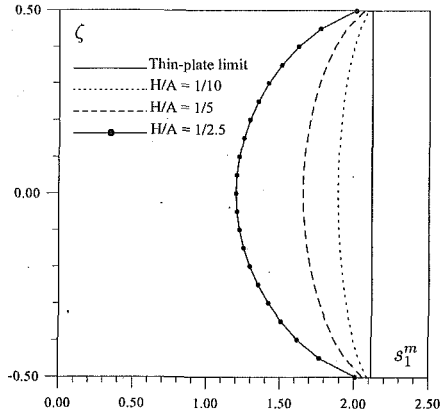


Fig. 2(*a*) Dimensionless side in-plane displacement s_1^m

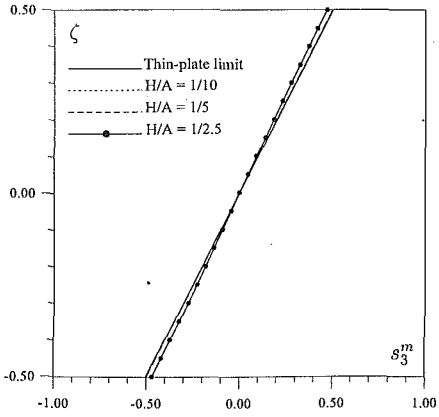


Fig. 2(*b*) Dimensionless central deflection s_3^m

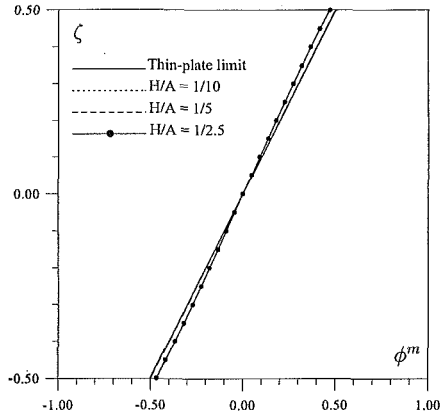


Fig. 2(*c*) Dimensionless central potential ϕ^m

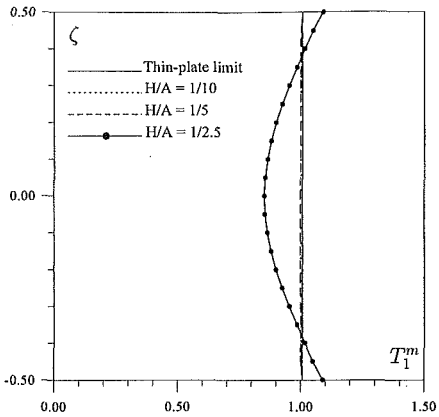


Fig. 2(*d*) Dimensionless central in-plane stress T_1^m

Fig. 2 Displacements, potential, and stress induced by ω_0

as $N_1, N_2 \to +\infty$. In the calculations it has been taken $N_1 = N_2 = 100$.

The results obtained are presented under the form of diagrams. The dimensionless variable ζ is reported on the ordinates axis. A dimensionless mechanical or electric quantity is reported on the abscissae axis. Four different lines are plotted in each diagram: three of them correspond to the values $H/A = 1/2.5$, $1/5$, $1/10$ of the thickness-to-side ratio of the plate; the fourth one represents the thin-plate limit. The material of the plate is the ceramic PZT-5H, presented in Table 1.

In Fig. 1 the case of a plate subjected to a uniform normal load of intensity p_0 ($p_0/2$ both on the upper and on the lower face: $p_3^+ = p_3^- = p_0/2$) is considered. In this case the plate has a bending behavior. The following dimensionless functions are plotted:

$$(a)\quad s_1^b(\zeta) := \frac{\pi^3 H^2 C}{3A^3 p_0}\, s_1\left(\frac{A}{2}, 0, H\zeta\right)$$

$$(b)\quad s_3^b(\zeta) := \frac{\pi^4 H^3 C}{3A^4 p_0}\, s_3(0, 0, H\zeta)$$

$$(c)\quad \phi^b(\zeta) := \frac{2\pi^2 H \epsilon C}{A^2 F p_0}\, \phi(0, 0, H\zeta)$$

$$(d)\quad T_1^b(\zeta) := \frac{\pi^2 H^2}{6(1-\xi)A^2 p_0}\, T_1(0, 0, H\zeta).$$

In Fig. 2 the case of a plate subjected to a uniform surface charge of density ω_0 on the upper face and of density $-\omega_0$ on the lower face ($\omega^+ = \omega_0$, $\omega^- = -\omega_0$) is considered. In this case the plate has a membrane behavior. The following dimensionless functions are plotted:

$$(a)\quad s_1^m(\zeta) := \frac{2\pi\epsilon C}{AF\omega_0}\, s_1\left(\frac{A}{2}, 0, H\zeta\right)$$

$$(b)\quad s_3^m(\zeta) := -\frac{\epsilon C}{HM\omega_0}\, s_3(0, 0, H\zeta)$$

$$(c)\quad \phi^m(\zeta) := \frac{\epsilon C}{HG\omega_0}\, \phi(0, 0, H\zeta)$$

$$(d)\quad T_1^m(\zeta) := -\frac{\epsilon}{F\xi\omega_0}\, T_1(0, 0, H\zeta).$$

It turns out that when the thickness-to-side ratio of the plate is less than or equal to $1/5$, the thin-plate limit provides results which agree with the exact ones within an error of 20 percent, for the evaluation of displacements, electric potential, and in-plane stress components; moreover, when that ratio is less than or equal to $1/10$, the maximum error reduces to 10 percent.

9 Conclusions

In this paper an exact three-dimensional solution for the problem of a simply supported transversely isotropic rectangular homogeneous piezoelectric plate is obtained in the framework of the linear theory of piezoelectricity.

The limit of this solution as the plate thickness aspect ratio approaches zero (thin-plate limit) is explicitly carried out. It represents a result to be agreed by a satisfactory theory of thin

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

piezoelectric plates. This target doesn't seem to be reached by existing theories. The results obtained in the present paper may represent a useful guide to develop a consistent theory of thin piezoelectric plates.

A numerical investigation in the case of a square plate is also performed in order to evaluate the influence of the thickness-to-side ratio on the three-dimensional solution of the plate problem. Electric and mechanical uniformly distributed loads are applied on the plate. This corresponds to a very frequent case in technical applications. It turns out that, when the thickness-to-side ratio of the plate is less than or equal to 1/5 (1/10), the thin-plate limit provides results which agree with the exact ones within an error of 20 percent (10 percent), for the evaluation of displacements, electric potential and in-plane stress components.

Acknowledgments

The authors express their gratitude to Prof. Elio Sacco for valuable comments on this paper.

The financial supports of the Italian National Research Council (CNR), of the Italian Ministry of University and Research (MURST) and of the Italian Space Agency (ASI) are gratefully acknowledged.

References

Ciarlet, P. G., and Destuynder, P., 1979, "A justification of the two-dimensional linear plate model," *J. Mécanique,* Vol. 18, pp. 315–344.

Gurtin, M. E., 1972, "The Linear Theory of Elasticity," *Handbuch der Physik,* S. Flügge, ed., Band VIa/2, Springer-Verlag, Berlin, pp. 129–131.

Eringen, A. C., and Maugin, G. A., 1990, *Electrodynamics of Continua,* Springer-Verlag, New York, pp. 239–257.

Ikeda, T., 1990, *Fundamentals of piezoelectricity,* Oxford University Press, Oxford, U.K., pp. 32–36.

Kiral, E., and Eringen, A. C., 1990, *Constitutive Equations of Nonlinear Electromagnetic-Elastic Crystals,* Springer-Verlag, Berlin, pp. 24–37.

Lekhnitskii, S. G., 1968, *Anisotropic plates,* (translated from the 2nd Russian ed. by S. W. Tsai and T. Cheron), Gordon and Breach, New York, pp. 318–392.

Levinson, M., 1985, "The simply supported rectangular plate: An exact, three dimensional, linear elasticity solution," *Journal of Elasticity,* Vol. 15, pp. 283–291.

Landau, L. D., and Lifshitz, E. M., 1981, *Electrodynamics of Continuous Media,* Course of Theoretical Physics, Vol. 8, Pergamon Press, Oxford, U.K., pp. 58–79.

Lee, Y. C., 1967, "Three-Dimensional Solution for Simply-Supported Thick Rectangular Plates," *Nuclear Engineering and Design,* Vol. 6, pp. 155–162.

Mason, W. P., 1981, "Piezoelectricity, its History and Applications," *J. Acoust. Soc. America,* Vol. 70, pp. 1561–1566.

Maugin, G. A., 1988, *Continuum Mechanics of Electromagnetic Solids,* North-Holland, Amsterdam, pp. 215–264.

Maugin, G. A., and Attou, D., 1990, "An asymptotic theory of thin piezoelectric plates," *Q. Jl. Mech. Appl. Math.,* Vol. 43, pp. 347–362.

Mindlin, R. D., 1972, "High frequency vibrations of piezoelectric crystal plates," *Int. J. Solids Structures,* Vol. 8, pp. 895–906.

Noor, A. K., and Scott Burton, W., 1990, "Three-Dimensional Solutions for Antisymmetrically Laminated Anisotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp. 182–188.

Pagano, N. J., 1970, "Exact Solutions for Rectangular Bidirectional Composites and Sandwich Plates," *J. Composite Materials,* Vol. 4, pp. 20–34.

Pan, E., 1991, "An exact solution for transversely isotropic, simply supported and layered rectangular plates," *Journal of Elasticity,* Vol. 25, pp. 101–116.

Parton, V. Z., 1984, "Modèles et Solutions Effectives en Elèctromagnèto-élasticité," *Proceedings IUTAM-IUPAP Symposium,* G. A. Maugin, ed., Elsevier Science Publishers B. V., Amsterdam, pp. 47–56.

Ray, M. C., Rao, K. M., and Samanta, B., 1992, "Exact Analysis of Coupled Electroelastic Behaviour of a Piezoelectric Plate under Cylindrical Bending," *Computers & Structures,* Vol. 45, pp. 667–677.

Smith, W. A., 1992, "The Key Design Principle for Piezoelectric Ceramic/Polymer Composites," *Proceedings Recent Advances in Adaptive and Sensory Materials and their Applications,* C. A. Rogers and R. C. Rogers, eds., Technomic Publishing Company, Lancaster, PA pp. 825–838.

Srinivas, S., and Rao, A. K., 1970, "Bending, vibration and buckling of simply supported thick orthotropic rectangular plates and laminates," *Int. J. Solids Structures,* Vol. 6, pp. 1463–1481.

Tiersten, H. F., 1969, *Linear Piezoelectric Plate Vibrations,* Plenum Press, New York, p. 58.

Timoshenko, S. P., and Woinowsky-Krieger, S., 1982, *Theory of plates and shells,* 23rd printing, Int. Student Ed., McGraw-Hill, New York, pp. 105–125.

Toshiba, 1983, *Piezoceramics,* Toshiba Ceramics Company Ltd, Tokyo, Japan, pp. 1–15.

Tzou, H. S., 1993, *Piezoelectric Shells,* Kluwer Academic Publishers, Dordrecht, pp. 13–71.

Varadan, V. V., Jeng, J.-H., and Varadan, V. K., 1987, "Form Invariant Constitutive Relations for Transversely Isotropic Piezoelectric Materials," *J. Acoust. Soc. Am.,* Vol. 82, pp. 337–341.

Vernitron, 1983, "Five Modern Piezoelectric ceramics," *Bulletin of Morgan Matroc Ltd,* Vernitron Piezoelectric Division, Bedford, OH, pp. 1–27.

Vlasov, B. F., 1957, "On One Case of Bending of Rectangular Thick Plates," *Vestnik Moskovskogo Universitieta,* (in Russian), No. 2, pp. 25–34.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A. P. S. Selvadurai
Professor and Chairman,
Department of Civil Engineering and
Applied Mechanics,
McGill University,
Montreal PQ H3A 2K6,
Canada
Mem. ASME

B. M. Singh
Research Fellow.

M. C. Au
Research Fellow.

Department of Civil Engineering,
Carleton University,
Ottawa, Ontario K1S 5B6
Canada

The Matrix-Fiber Crack in an Elastic Solid

In the presence of strong interface adhesion, the fracture of an embedded fiber can also result in the cracking of the surrounding matrix. While the orientation of such matrix cracks can be varied, the flat penny-shaped crack represents a critical crack orientation which is of particular interest to the study of the micromechanics of fracture processes in fiber-reinforced solids. This paper considers the axisymmetric problem of the uniform straining of a composite elastic solid which contains a penny-shaped crack occupying both the fiber and matrix regions. The isolated cracked fiber-matrix crack interaction is formulated as a mixed boundary value problem related to a two-domain half-space region. The resulting integral equations are solved in a numerical fashion to evaluate the stress intensity factor at the boundary of the penny-shaped crack. The numerical results presented in the paper illustrate the influence of the elasticity mismatch between the fiber and the matrix on the stress intensity factor at the crack-tip located in the matrix. The numerical results are presented for typical fiber-reinforced composites consisting of epoxy and ceramic matrices reinforced with silicon, glass, and kevlar fibers.

Introduction

The interaction of cracks and fibers is an important consideration in the modeling of micromechanics of damage in fiber-reinforced solids. Such damage can take a variety of forms depending upon the constitutive properties of the matrix, the reinforcing fibers and the conditions at the fiber-matrix interface (Backlund, 1981). With predominantly brittle elastic fiber-reinforced solids, the damage takes the form of matrix cracking, fiber-matrix interface delamination and fiber fracture (see, e.g., Dvorak, 1991). For example, matrix cracking in the presence of intact fibres results in the process of flaw and crack bridging which has been investigated by a number of researchers including Aveston et al. (1971), Kelly (1970), Sih (1979), and Selvadurai (1981). In recent years a number of investigators including Marshall et al. (1985), Budiansky et al. (1986), McCartney (1987), Stang (1987), and Rose (1987) have extended and improved the flaw bridging model to include various categories of linear and nonlinear flaw bridging processes. While matrix cracking and bridging usually results from loading of the composites in the direction of the fibers, interface delaminations can result from thermal loadings and loadings normal to the fiber directions and in-plane shearing. In many instances, interface delamination is usually treated as a two-dimensional plane-strain problem which considers the cross section of a unidirectional fiber-reinforced composite normal to the fiber direction. In this case, the classical solutions related to inclusion arrays in a matrix with delaminations at the fiber-matrix interface can be used to examine the problem. In addition to the classical studies by England (1965) and Dundurs and Mura (1964), references to developments in this area are given by Sendeckyi (1974), Chen and Sih (1985), and Mura (1988, 1989). These studies have also been extended to include cases where the interface debonded region is subjected to antiplane shear (Sendeckyi, 1974; Atkinson, 1979; Karihaloo and Viswanathan, 1985). The complete interface delamination problem is three-dimensional where regions of delamination can occur in a random fashion. Özbek and Erdogan (1969) and Erdogan and Özbek (1969) (see also Cherepanov, 1979) have examined the axisymmetric problem of an infinite elastic fiber embedded into an infinite elastic matrix where interface debonding occurs over a cylindrical region. These problems simulate the influence of thermal loadings of the debonded composite region. A majority of interface delamination problems involving composite regions assume loss of contact at the delaminated region. Processes such as interface frictional locking, slip, closure, etc., have an important influence on the mechanical behavior of the interface crack (Kelly, 1970; Comninou and Schmueser, 1979).

When the fiber-matrix interface exhibits strong adhesive toughness characteristics, fiber fracture usually results in the development of matrix fracture. This mode of fiber-matrix failure has been observed in fragmentation tests conducted to determined the adhesive toughness of fiber-matrix interfaces. Examples of this mode of failure are given by Chamis (1974) and van den Berg (1990), Busschen and Selvadurai (1995), and Selvadurai and Busschen (1995). The fracture topography of such combined fiber-matrix failure is usually quite complex. The matrix crack resulting in the presence of strong interface adhesion can be conoidal or a combination of conoidal and flat penny-shaped cracks. The interaction of cracks and fibers have been investigated by a number of researchers. Pacella and Erdogan (1974) and Narayan and Erdogan (1975) have examined the axisymmetric elastic problems where a penny-shaped matrix crack is surrounded by regular arrays of elastic fibers which are located in the vicinity of the crack. Dhaliwal et al. (1979) have examined the problem of the internal pressurization of a penny-shaped crack which is located at the interior of an embedded elastic fiber. These studies have also been extended to include the mechanics of an external matrix crack surrounding an elastic fiber (Singh et al. 1979). Recently, Wijeywickrema et al. (1990) have examined the interaction between an elastic fiber and an annular crack located in the matrix surrounding the fiber. Of related interest is the class of problems which deal with the interaction of inclusions and cracks. Taya and Mura (1981) and Taya and Chou (1983) have examined the problem of a penny-shaped crack which is located at the extremity of an elongated ellipsoidal elastic inclusion embedded in an elastic medium. Selvadurai (1991a) has utilized the boundary integral equation technique to examine the behavior of penny-shaped cracks

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, June 25, 1993; final revision, Jan. 17, 1996. Associate Technical Editor: D. M. Parks.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

which are located at the boundary of cylindrical elastic inclusions.

In this paper, we restrict attention to matrix cracking which accompanies the complete fracture of an isolated fiber. It is assumed that this fracture configuration results in a defect which can be modeled as a penny-shaped crack located in a single fibre-matrix composite solid. The treatment of the matrix as a domain of infinite extent is of course a highly idealized approximation to an actual situation of local damage in a fiber-reinforced composite. A more rigorous and realistic treatment of the cracked fiber problem related to a composite material should also take into consideration the influence of neighboring uncracked fibers, random locations of matrix fiber cracks, off-axis location of the matrix crack, and possible obliquity of the penny-shaped crack. The analysis of the single matrix-fiber crack presented in the paper is applicable to situations where low fiber volume concentrations results in fiber spacings which are large in comparison with the diameter of the fibers. The resulting idealization deals with the problem of an isolated fiber which is located in an isotropic elastic matrix of infinite extent. The composite region is weakened by a penny-shaped crack which occupies both the matrix and the single fiber regions (Fig. 1). The composite region is subjected to a state of uniaxial strain along the fiber direction. Owing to the symmetry of matrix-fiber crack region about the plane $z = 0$, the analysis of the problem can be formulated as a mixed boundary value problem related to a two-domain half-space region. The resulting system of Fredholm integral equations of the second kind are solved in a numerical fashion to generate the stress intensity factor at the boundary of the penny-shaped crack.

The Matrix-Fiber Crack Problem

An isotropic elastic matrix of infinite extent contains an elastic fibre which is perfectly bonded to it. The composite region is weakened by a penny-shaped crack which is generated by the fracture of the fiber and that of the matrix. The radius of the fiber and that of the matrix crack are denoted by a and b, respectively. The composite region is subjected to a uniform axial strain ϵ_0 at regions remote from the penny-shaped crack. The solution of the crack problem follows the conventional approach of combining two solutions; the first solution corresponds to the uniform straining of an uncracked composite region and a corrective solution which renders the region occupying the crack, free of surface tractions.

For the solution of the uncracked composite region we consider Lamé-type stress and displacement fields in the matrix (suffix m) and fiber (suffix f) regions which take the following forms (Timoshenko and Goodier, 1970):

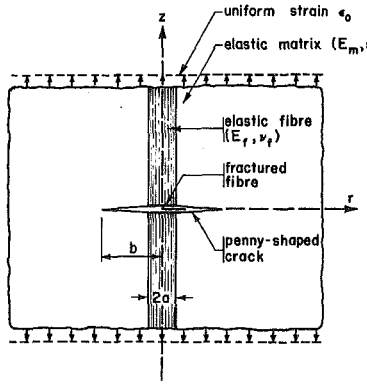


Fig. 1 The matrix-fiber crack in an elastic solid

$$\sigma_{rr}^{(m)}(r, z) = \frac{A_m}{r^2} + 2B_m \tag{1a}$$

$$\sigma_{\theta\theta}^{(m)}(r, z) = -\frac{A_m}{r^2} + 2B_m \tag{1b}$$

$$u_r^{(m)}(r, z) = \frac{r}{E_m}\left\{2B_m - \frac{A_m}{r^2} - \nu_m\left[2B_m + 4\nu_m B_m + E_m\epsilon_0 + \frac{A_m}{r^2}\right]\right\} \tag{1c}$$

and

$$\sigma_{rr}^{(f)}(r, z) = \sigma_{\theta\theta}^{(f)}(r, z) = 2C_f \tag{2a}$$

$$u_r^{(f)}(r, z) = \frac{r}{E_f}\{2C_f - \nu_f[2C_f + 4\nu_f C_f + E_f\epsilon_0]\} \tag{2b}$$

where E_i, ν_i are the elastic constants of the fiber ($i = f$) and matrix ($i = m$) regions and A_m, B_m, and C_f are constants which are to be determined by making use of the traction and displacement continuity conditions at the fiber-matrix boundary and the far-field stress state: i.e.,

$$\sigma_{rr}^{(f)}(a, z) = \sigma_{rr}^{(m)}(a, z) \tag{3a}$$

$$u_r^{(f)}(a, z) = u_r^{(m)}(a, z) \tag{3b}$$

$$\sigma_{rr}^{(m)}(r, z) \to 0 \text{ as } (r^2 + z^2)^{1/2} \to \infty. \tag{3c}$$

Considering the above results it can be shown that the axial stresses present in the fiber and matrix regions in the presence of complete interface bonding and uniaxial straining in the z-direction take the forms

$$\sigma_{zz}^{(f)}(r, z) = \Omega\sigma_0 \tag{4a}$$

$$\sigma_{zz}^{(m)}(r, z) = \sigma_0 \tag{4b}$$

where

$$\sigma_0 = E_m\epsilon_0 \tag{5a}$$

$$\Omega = \frac{\mu_f}{\mu_m}\left\{\frac{(1 + \nu_f)}{(1 + \nu_m)} + \frac{2\nu_f(\nu_f - \nu_m)}{(1 + \nu_m)\left[1 - 2\nu_f + \frac{\mu_f}{\mu_m}\right]}\right\} \tag{5b}$$

and μ_i are the linear elastic shear moduli of the fiber ($i = f$) and matrix ($i = m$) regions, respectively.

In the corrective solution, the composite region is subjected to equal and opposite axial stresses defined by (4a) and (4b) in the region corresponding to the crack. Considering the symmetry of the cracked fiber-matrix region about the plane $z = 0$, the corrective solution can be formulated as a mixed boundary value problem associated with a two-domain half-space region occupying the region $z \geq 0$. The boundary and interface conditions corresponding to the corrective solution are as follows:

$$u_z^{(f)}(a, z) = u_z^{(m)}(a, z); \quad 0 \leq z < \infty \tag{6}$$

$$u_r^{(f)}(a, z) = u_r^{(m)}(a, z); \quad 0 \leq z < \infty \tag{7}$$

$$\sigma_{rr}^{(f)}(a, z) = \sigma_{rr}^{(m)}(a, z); \quad 0 \leq z < \infty \tag{8}$$

$$\sigma_{rz}^{(f)}(a, z) = \sigma_{rz}^{(m)}(a, z); \quad 0 \leq z < \infty \tag{9}$$

$$\sigma_{rz}^{(f)}(r, 0) = 0; \quad 0 < r < a \tag{10}$$

$$\sigma_{rz}^{(m)}(r, 0) = 0; \quad a < r < \infty \tag{11}$$

$$\sigma_{zz}^{(f)}(r, 0) = -\Omega\sigma_0; \quad 0 < r < a \tag{12}$$

$$\sigma_{zz}^{(m)}(r, 0) = -\sigma_0; \quad a < r < b \tag{13}$$

$$u_z^{(m)}(r, 0) = 0; \quad b \leq r < \infty. \tag{14}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

In addition, the displacement and stress fields in the two-domain half-space region should satisfy the regularity conditions

$$u_\alpha^{(i)},\quad \sigma_{\alpha\beta}^{(i)} \to 0 \text{ as } (r^2 + z^2)^{1/2} \to \infty;\quad i = f, m.$$

The solution of the mixed boundary value problem defined by (6) to (14) yields the appropriate singular stress field at the boundary of the penny-shaped crack which can be used to evaluate the stress intensity factor at the crack-tip.

The Governing Equations

For the solution of the axisymmetric problem governing the corrective stress state, it is convenient to employ the strain potential approach proposed by Love (1944). It can be shown that in the absence of body forces the displacement and stress fields can be expressed in terms of a function $\chi^{(i)}(r, z)$, $(i = f, m)$ which satisfies

$$\nabla^2\nabla^2\chi^{(i)}(r, z) = 0 \tag{15}$$

where ∇^2 is Laplace's operator referred to the cylindrical polar coordinate system: i.e.,

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}. \tag{16}$$

The expressions for the displacements and stresses take the form

$$2\mu_i u_r^{(i)}(r, z) = -\frac{\partial^2\chi^{(i)}}{\partial r\partial z} \tag{17}$$

$$2\mu_i u_z^{(i)}(r, z) = \left[(1 - 2\nu_i)\nabla^2 + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right]\chi^{(i)} \tag{18}$$

and

$$\sigma_{rr}^{(i)}(r, z) = \frac{\partial}{\partial z}\left\{\nu_i\nabla^2 - \frac{\partial^2}{\partial r^2}\right\}\chi^{(i)} \tag{19}$$

$$\sigma_{\theta\theta}^{(i)}(r, z) = \frac{\partial}{\partial z}\left\{\nu_i\nabla^2 - \frac{1}{r}\frac{\partial}{\partial r}\right\}\chi^{(i)} \tag{20}$$

$$\sigma_{zz}^{(i)}(r, z) = \frac{\partial}{\partial z}\left\{(2 - \nu_i)\nabla^2 - \frac{\partial^2}{\partial z^2}\right\}\chi^{(i)} \tag{21}$$

$$\sigma_{rz}^{(i)}(r, z) = \frac{\partial}{\partial r}\left\{(1 - \nu_i)\nabla^2 - \frac{\partial^2}{\partial z^2}\right\}\chi^{(i)}, \tag{22}$$

respectively.

To formulate the mixed boundary value problem defined by (6) to (14) we seek solutions of (15) applicable to the fiber and matrix regions, respectively. A suitable biharmonic function applicable to the fiber region $(r\epsilon(0, a);\ z\epsilon(0, \infty))$ can be taken as (Sneddon, 1951)

$$\chi^{(f)}(r, z) = -2\mu_f\int_0^\infty s^{-3}F_1(s)[2\nu_f + sz]e^{-sz}J_0(sr)ds$$

$$- 2\mu_f\int_0^\infty s^{-2}[\{C(s) + 4(1 - \nu_f)D(s)\}I_0(sr) - srD(s)I_1(sr)]\sin(sz)ds \tag{23}$$

where $F_1(s)$, $C(s)$, and $D(s)$ are unknown functions which are to be determined and $J_p(\alpha)$ and $I_p(\alpha)$ denote, respectively, Bessel functions of the first kind and modified Bessel functions of the second kind of order $p(\geq 0)$. The components of the displacement vector and the stress tensor for the fiber region can be obtained with the aid of Eqs. (17) to (22). A biharmonic function $\chi^{(m)}(r, z)$ applicable to the matrix region $(r\epsilon(a, \infty);\ z\epsilon(0, \infty))$ takes the form

$$\chi^{(m)}(r, z) = -2\mu_m\int_0^\infty s^{-3}F_2(s)[2\nu_m + sz]e^{-sz}J_0(sr)ds$$

$$- 2\mu_m\int_0^\infty s^{-2}[\{A(s) - 4(1 - \nu_m)B(s)\}K_0(sr) - srK_1(sr)B(s)]\sin(sz)ds \tag{24}$$

where $A(s)$, $B(s)$, and $F_2(s)$ are unknown functions and $K_p(\alpha)$ are modified Bessel functions of the first kind of order $p(\geq 0)$. The components of the displacement vector and the stress tensor for the matrix region can be obtained by making use of (24) and expressions (17) to (22).

Integral Equations Governing the Matrix-Fiber Crack Problems

The integral expressions for $u_r^{(i)}$, $u_z^{(i)}$, $\sigma_{rr}^{(i)}$, $\sigma_{zz}^{(i)}$, and $\sigma_{rz}^{(i)}$ $(i = f, m)$ obtained from (23) and (24) can be used to formulate the integral equations governing the matrix-fiber crack problem. Considering the matrix region we observe the boundary condition (11) is identically satisfied and the boundary conditions (13) and (14) yield the following integral equations:

$$\int_0^\infty F_2(s)J_0(sr)ds + \int_0^\infty s\{A(s)K_0(sr) + [2\nu_m K_0(sr) - srK_1(sr)]B(s)\}ds = \frac{\sigma_0}{2\mu_m};\quad a < r < b \tag{25}$$

$$\int_0^\infty s^{-1}F_2(s)J_0(sr)ds = 0;\quad b \leq r < \infty. \tag{26}$$

Consider a trial solution of $F_2(s)$ of the form

$$F_2(s) = s\left[\int_a^b h(t)\cos(st)dt + \int_b^\infty g(t)\cos(st)dt\right] \tag{27}$$

where $h(t)$ and $g(t)$ are arbitrary functions and $g(\infty) = 0$. Substituting the above representation in (26) we obtain an Abel-type integral equation

$$\int_b^r \frac{g(t)dt}{(r^2 - t^2)^{1/2}} + \int_a^b \frac{h(t)dt}{(r^2 - t^2)^{1/2}} = 0;\quad b \leq r < \infty. \tag{28}$$

The solution of (28) can be written as

$$g(t) = -\frac{2}{\pi}\frac{t}{(t^2 - b^2)^{1/2}}\int_a^b \frac{h(u)(b^2 - u^2)^{1/2}}{(t^2 - u^2)}du;\quad b \leq t < \infty. \tag{29}$$

Similarly, using the representation (27), the integral Eq. (25) can be reduced to the form

$$h(t) - \frac{2}{\pi}(b^2 - t^2)^{1/2}\int_b^\infty \frac{ug(u)du}{(u^2 - t^2)(u^2 - b^2)^{1/2}} + \frac{2}{\pi}\int_0^\infty s\{A(s)L_1(s, t) + [2\nu_m L_1(s, t) - sL_2(s, t)]B(s)\}ds = \frac{\sigma_0(b^2 - t^2)^{1/2}}{\pi\mu_m};\quad a < t < b \tag{30}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$L_1(s, t) = \int_t^b \frac{rK_0(sr)dr}{(r^2 - t^2)^{1/2}} \tag{31a}$$

$$L_2(s, t) = \int_t^b \frac{r^2K_1(sr)dr}{(r^2 - t^2)^{1/2}}. \tag{31b}$$

Considering the fiber-matrix interface continuity conditions (6) to (9), we obtain via a Fourier inversion the following:

$$saI_1(sa)D(s) - I_0(sa)C(s) - saK_1(sa)B(s) + K_0(sa)A(s)$$
$$= \frac{2}{\pi}\int_0^\infty u^{-1}[\{2(1 - \nu_m)f_1 + uf_2\}F_2(u) - \{2(1 - \nu_f)f_1 + uf_2\}F_1(u)]J_0(ua)du = X_1(s); \quad 0 \le s < \infty \tag{32}$$

$$I_1(sa)C(s) + \{4(1 - \nu_f)I_1(sa) - saI_0(sa)\}D(s) + K_1(sa)A(s) - \{4(1 - \nu_m)K_1(sa) + saK_0(sa)\}B(s)$$
$$= \frac{2}{\pi}\int_0^\infty u^{-1}[\{(2\nu_m - 1)f_3 + uf_4\}F_2(u) - \{(2\nu_f - 1)f_3 + uf_4\}F_1(u)]J_1(ua)du = X_2(s); \quad 0 \le s < \infty \tag{33}$$

$$\mu_m\{[K_1(sa) + saK_0(sa)]A(s) - [\{4(1 - \nu_m) + s^2a^2\}K_1(sa) + (3 - 2\nu_m)saK_0(sa)]B(s)\} + \mu_f\{[I_1(sa) - saI_0(sa)]C(s) + [\{4(1 - \nu_f) + s^2a^2\}I_1(sa) - (3 - 2\nu_f)saI_0(sa)]D(s)\}$$
$$= \frac{2}{\pi}\mu_f a\int_0^\infty \left[\{-f_3 + uf_4\}J_0(ua) + \{(1 - 2\nu_f)f_3 - uf_4\}\frac{J_1(ua)}{ua}\right]F_1(u)du - \frac{2}{\pi}\mu_m a\int_0^\infty \left[\{-f_3 + uf_4\}J_0(ua) + \{(1 - 2\nu_m)f_3 - uf_4\}\frac{J_1(ua)}{ua}\right]F_2(u)du = \mu_f X_3(s); \quad 0 \le s < \infty \tag{34}$$

$$\mu_m\{A(s)K_1(sa) - [2(1 - \nu_m)K_1(sa) + saK_0(sa)]B(s)\} + \mu_f\{C(s)I_1(sa) + [2(1 - \nu_f)I_1(sa) - saI_0(sa)]D(s)\}$$
$$= \frac{2}{\pi}\frac{\mu_m}{s}\int_0^\infty uF_2(u)f_2J_1(ua)du - \frac{2}{\pi}\frac{\mu_f}{s}\int_0^\infty uF_1(u)f_2J_1(ua)du = \frac{\mu_f}{s}X_4(s); \quad 0 \le s \le \infty \tag{35}$$

where

$$f_1 = \int_0^\infty e^{-uz}\sin(sz)dz = \frac{s}{(s^2 + u^2)} \tag{36}$$

$$f_2 = \int_0^\infty ze^{-uz}\sin(sz)dz = \frac{2su}{(s^2 + u^2)^2} \tag{37}$$

$$f_3 = \int_0^\infty e^{-uz}\cos(sz)dz = \frac{u}{(s^2 + u^2)} \tag{38}$$

$$f_4 = \int_0^\infty ze^{-uz}\cos(sz)dz = \frac{(u^2 - s^2)}{(u^2 + s^2)^2}. \tag{39}$$

In the sequel, it is convenient to adopt the following convention:

$$I_0(sa) = I_0; \quad I_1(sa) = I_1 \tag{40}$$

$$K_0(sa) = K_0; \quad K_1(sa) = K_1. \tag{41}$$

The Eqs. (32) to (35) can now be solved to evaluate the arbitrary functions $A(s)$, $B(s)$, $C(s)$, and $D(s)$ in terms of $X_i(s)$ $(i = 1, 2, 3, 4)$ or consequently in terms of $F_1(s)$ and $F_2(s)$. These expressions are given in Appendix A.

Considering the boundary conditions (10) and (12) it can be shown that the boundary condition (10) is identically satisfied by $\sigma_{rz}^{(f)}$ and that the expression for $\sigma_{zz}^{(f)}$ yields the following integral equation:

$$\int_0^\infty F_1(s)J_0(sr)ds + \int_0^\infty s\{C(s)I_0(sr) - [2\nu_fI_0(sr) + srI_1(sr)]D(s)\}ds = \frac{\Omega\sigma_0}{2\mu_f}; \quad 0 < r < a. \tag{42}$$

Consider the representation

$$F_1(s) = s\int_0^a m(t)\sin(st)dt. \tag{43}$$

We note that the Eq. (42) can be rewritten in the form

$$\frac{\partial}{\partial r}\left\{r\int_0^\infty \frac{F_1(s)}{s}J_1(sr)ds\right\} + r\int_0^\infty s\{C(s) - 2\nu_fD(s)\}I_0(sr)ds - r^2\int_0^\infty s^2I_1(sr)D(s)ds = \frac{\Omega\sigma_0}{2\mu_f}; \quad 0 < r < a. \tag{44}$$

Using (43), (44) can be expressed in the form

$$\frac{\partial}{\partial r}\int_0^r \frac{tm(t)dt}{(r^2 - t^2)^{1/2}} + r\int_0^\infty s\{C(s) - 2\nu_fD(s)\}I_0(sr) - r^2\int_0^\infty s^2I_1(sr)D(s)ds = \frac{\Omega\sigma_0 r}{2\mu_f}; \quad 0 < r < a. \tag{45}$$

Making use of the following results

$$\int_0^t \frac{rI_0(sr)ds}{(t^2 - r^2)^{1/2}} = \frac{\sinh(st)}{s} \tag{46}$$

$$\int_0^t \frac{r^2I_1(sr)dr}{(t^2 - r^2)^{1/2}} = \frac{st\cosh(st) - \sinh(st)}{s^2}, \tag{47}$$

the integral Eq. (45) can be expressed in the form

$$m(t) + \frac{2}{\pi}\int_0^\infty [X_1(s)R_1(s, t) + X_2(s)R_2(s, t) + X_3(s)R_3(s, t) + X_4(s)R_4(s, t)]ds = \frac{\Omega\sigma_0 t}{\pi\mu_f}; \quad 0 < t < a \tag{48}$$

where

$$\eta R_1(s, t) = \sinh(st)(a_7a_{13} - a_8a_{12}) - (a_{13}I_0 + a_8I_1)\{\sinh(st)(2\nu_f - 1) + st\cosh(st)\} \tag{49}$$

$$\eta R_2(s, t) = \sinh(st)(a_7a_{14} - a_9a_{12}) - (a_{14}I_0 + a_9I_1)\{\sinh(st)(2\nu_f - 1) + st\cosh(st)\} \tag{50}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\eta R_3(s, t) = \sinh(st)(a_7a_{15} - a_{10}a_{12}) - (a_{15}I_0 + a_{10}I_1)\{\sinh(st)(2\nu_f - 1) + st\cosh(st)\} \quad (51)$$

$$\eta R_4(s, t) = \sinh(st)(a_7a_{16} - a_{11}a_{12}) - (a_{16}I_0 + a_{11}I_1)\{\sinh(st)(2\nu_f - 1) + st\cosh(st)\}. \quad (52)$$

Considering the integrals of Eqs. (32) to (35) containing $F_1(u)$, the integral expressions (36) to (39) and the representation (43) it can be shown that

$$-\frac{2}{\pi}\int_0^\infty \{2(1-\nu_f)f_1 + uf_2\}\frac{F_1(u)}{u}J_0(ua)du = -\frac{2}{\pi}\int_0^a [(3-2\nu_f)\sinh(su)K_0(sa) - saK_1(sa)\sinh(su) + su\cosh(su)K_0(sa)]m(u)du; \quad 0 < s < \infty \quad (53)$$

$$-\frac{2}{\pi}\int_0^\infty \{(2\nu_f - 1)f_3 + uf_4\}\frac{F_1(u)}{u}J_1(ua)du = -\frac{2}{\pi}\int_0^a [(2\nu_f - 1)K_1(sa)\sinh(su) - \{saK_0(sa)\sinh(su) - suK_1(sa)\cosh(su)\}]m(u)du; \quad 0 < s < \infty \quad (54)$$

$$\frac{2}{\pi}a\mu_f\int_0^\infty \left\{(-f_3 + uf_4)J_0(ua) + [(1-2\nu_f)f_3 - uf_4]\frac{J_1(ua)}{ua}\right\}F_1(u)du = \frac{2}{\pi}\mu_f\int_0^a \{sa[sa\sinh(su)K_1(sa) - su\cosh(su)K_0(sa)] + (1-2\nu_f)K_1(sa)\sinh(su) - suK_1(sa)\cosh(su)\}m(u)du; \quad 0 < s < \infty \quad (55)$$

$$-\frac{2}{\pi}\frac{\mu_f}{s}\int_0^\infty uF_1(u)f_2J_1(ua)du = -\frac{2\mu_f}{\pi}\int_0^a [\sinh(su)K_1(sa) + su\cosh(su)K_1(sa) - sa\sinh(su)K_0(sa)]m(u)du; \quad 0 < s < \infty. \quad (56)$$

In a similar manner, the integrals of Eqs. (32) to (35) containing $F_2(u)$ can be expressed in terms of the functions $h(u)$ and $g(u)$ (defined by (27)) by the following expressions:

$$\frac{2}{\pi}\int_0^\infty u^{-1}[\{2(1-\nu_m)f_1 + uf_2\}F_2(u)J_0(ua)du = [(3-2\nu_m)I_0(sa) + saI_1(sa)]\times\left[\int_a^b h(u)e^{-su}du + \int_b^\infty g(u)e^{-su}du\right] - sI_0(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right]; \quad 0 < s < \infty \quad (57)$$

$$\frac{2}{\pi}\int_0^\infty u^{-1}[\{(2\nu_m - 1)f_3 + uf_4\}F_2(u)]J_1(ua)du = -\left\{\frac{sa}{2}[I_0(sa) + I_2(sa)] + 2\nu_mI_1(sa)\right\}\times\left[\int_a^b h(u)e^{-su}du + \int_b^\infty g(u)e^{-su}du\right] + sI_1(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right]; \quad 0 < s < \infty \quad (58)$$

$$-\frac{2}{\pi}a\mu_m\int_0^\infty \left[(-f_3 + uf_4)J_0(ua) + \{(1-2\nu_m)f_3 - uf_4\}\frac{J_1(ua)}{ua}\right]F_2(u)du = \mu_m\left[\int_a^b h(t)e^{-st}\left[sa\left\{\frac{1}{2}I_0(sa) + saI_1(sa) - stI_0(sa)\right\} + I_1(sa)(st - 2\nu_m) - \frac{sa}{2}I_2(sa)\right]dt\right] + \mu_m\left[\int_b^\infty g(t)e^{-st}\left[sa\left\{\frac{1}{2}I_0(sa) + saI_1(sa) - stI_0(sa)\right\} + I_1(sa)(st - 2\nu_m) - \frac{sa}{2}I_2(sa)\right]dt\right]; \quad 0 < s < \infty \quad (59)$$

$$\frac{2}{\pi}\mu_m\int_0^\infty uF_2(u)f_2J_1(ua)du = -s\mu_m\left[2I_1(sa) + \frac{sa}{2}(I_0(sa) + I_2(sa))\right]\times\left[\int_a^b h(u)e^{-su}du + \int_b^\infty e^{-su}g(u)du\right] + s^2\mu_mI_1(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right]; \quad 0 < s < \infty. \quad (60)$$

Making use of the expressions (53) to (60) it is possible to obtain expressions for $X_i(s)$ $(i = 1, 2, 3, 4)$ in terms of the functions $h(s)$, $g(s)$, and $m(s)$. These expressions are given in Appendix A. Using the results given in Eqs. (A24) to (A27), the integral Eq. (30) can be expressed in the form

$$H(t) - \frac{2}{\pi}(b^2 - t^2)^{1/2}\int_b^\infty \frac{uG(u)du}{(u^2 - t^2)(u^2 - b^2)^{1/2}} + \frac{2}{\pi}\int_a^b H(u)S_1(u, t)du + \frac{2}{\pi}\int_b^\infty G(u)S_1(u, t)du + \frac{4}{\pi^2}\int_0^a M(u)S_2(u, t)du = (b^2 - t^2)^{1/2}; \quad a < t < b \quad (61)$$

where

$$H(t) = \frac{\pi\mu_m}{\sigma_0}h(t) \quad (62)$$

$$G(t) = \frac{\pi\mu_m}{\sigma_0}g(t) \quad (63)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$M(t) = \frac{\pi\mu_m}{\sigma_0} m(t), \tag{64}$$

and the kernel functions S_1 and S_2 are given by

$$S_1(u, t) = \int_0^\infty e^{-su}[\{(3 - 2\nu_m)I_0(sa) + saI_1(sa) - suI_0(sa)\}P_1(s, t) - \left\{\frac{sa}{2}(I_0(sa) + I_2(sa)) + 2\nu_m I_1(sa) - suI_1(sa)\right\}P_2(s, t) + \Gamma\left\{sa\left[\frac{1}{2}I_0(sa) + saI_1(sa) - suI_0(sa)\right] + I_1(sa)(su - 2\nu_m) - \frac{sa}{2}I_2(sa)\right\}P_3(s, t) - \Gamma\left\{2sI_1(sa) + \frac{s^2a}{2}[I_0(sa) + I_2(sa)] - s^2uI_1(sa)\right\}P_4(s, t)\Big]ds \tag{65}$$

$$S_2(u, t) = -\int_0^\infty [\{(3 - 2\nu_f)\sinh(su)K_0(sa) - saK_1(sa)\sinh(su) + su\cosh(su)K_0(sa)\}P_1(s, t) + \{(2\nu_f - 1)K_1(sa)\sinh(su) - [sa\sinh(su)K_0(sa) - su\cosh(su)K_1(sa)]\}P_2(s, t) - \{sa[sa\sinh(su)K_1(sa) - su\cosh(su)K_0(sa)] + (1 - 2\nu_f)K_1(sa)\sinh(su) - suK_1(sa)\cosh(su)\}P_3(s, t) + s\{\sinh(su)K_1(sa) + su\cosh(su)K_1(sa) - sa\sinh(su)K_0(sa)\}P_4(s, t)]ds. \tag{66}$$

The expressions for the functions $P_i(s, t)$ $(i = 1, 2, 3, 4)$ occurring in Eqs. (65) and (66) are also given in Appendix A.

Using a similar procedure, the integral Eq. (48) can be reduced to the form

$$M(t) + \frac{2}{\pi}\int_a^b H(u)S_3(u, t)du + \frac{2}{\pi}\int_b^\infty G(u)S_3(u, t)du + \frac{4}{\pi^2}\int_0^a M(u)S_4(u, t)du = t\Gamma\Omega; \quad 0 < t < a \tag{67}$$

where

$$S_3(u, t) = \int_0^\infty e^{-su}\Big[\{(3 - 2\nu_m)I_0(sa) + saI_1(sa) - suI_0(sa)\}R_1(s, t) - \left\{\frac{sa}{2}[I_0(sa) + I_2(sa)] + 2\nu_m I_1(sa) - suI_1(sa)\right\}R_2(s, t) + \Gamma\left\{sa\left[\frac{1}{2}I_0(sa) + saI_1(sa) - suI_0(sa)\right] + I_1(sa)[su - 2\nu_m] - \frac{sa}{2}I_2(sa)\right\}R_3(s, t) - \Gamma\left\{2sI_1(sa) + \frac{s^2a}{2}[I_0(sa) + I_2(sa)] - s^2uI_1(sa)\right\}R_4(s, t)\Big]ds. \tag{68}$$

$$S_4(u, t) = -\int_0^\infty [\{(3 - 2\nu_f)\sinh(su)K_0(sa) - saK_1(sa)\sinh(su) + su\cosh(su)K_0(sa)\}R_1(s, t) + \{(2\nu_f - 1)K_1(sa)\sinh(su) - [sa\sinh(su)K_0(sa) - su\cosh(su)K_1(sa)]\}R_2(s, t) - \{sa[sa\sinh(su)K_1(sa) - su\cosh(su)K_0(sa)] + (1 - 2\nu_f)K_1(sa)\sinh(su) - suK_1(sa)\cosh(su)\}R_3(s, t) + s\{\sinh(su)K_1(sa) + su\cosh(su)K_1(sa) - sa\sinh(su)K_0(sa)\}R_4(s, t)]ds \tag{69}$$

where the expressions for the functions R_i $(i = 1, 2, 3, 4)$ occurring in the kernel functions $S_3(u, t)$ and $S_4(u, t)$ are given in Appendix A.

Finally, by using the transformations given by (62) and (63), Eq. (28) can be rewritten in the form

$$G(t) + \frac{2}{\pi}\frac{t}{(t^2 - b^2)^{1/2}}\int_a^b \frac{H(u)(b^2 - u^2)^{1/2}}{(t^2 - u^2)}du = 0; \quad b \le t < \infty. \tag{70}$$

Equations (61), (67), and (70) represent the system of coupled integral equations for the unknown functions $H(t)$, $G(t)$, and $M(t)$ which govern the mixed boundary value problem for the penny-shaped matrix-fiber crack for an elastic solid which is subjected to uniform straining. It is evident that owing to the complexity of the kernel functions $S_i(u, t)$ $(i = 1, 2, 3, 4)$ occurring in these equations recourse must be made to solve these coupled integral equations in a numerical fashion. The details of the numerical scheme are given in Appendix B.

The Stress Intensity Factor

A result of primary importance to the examination of brittle elastic fracture processes of the matrix region is the stress intensity factor at the boundary, $r = b$, of the matrix-fiber crack. A critical value of the stress intensity factor is an indicator of the fracture toughness of the matrix region. In the case of the penny-shaped matrix-fiber crack, the magnitude of the stress intensity factor will be influenced by the elasticity mismatch between the fiber and matrix regions. Developing an expression for $\sigma_{zz}^{(m)}$ we have

$$\frac{\sigma_{zz}^{(m)}(r, 0)}{2\mu_m} = -\int_0^\infty F_2(s)J_0(sr)ds - \int_0^\infty s\{A(s)K_0(sr) + [2\nu_m K_0(sr) - srK_1(sr)]B(s)\}ds; \quad a < r < b. \tag{71}$$

The stress intensity factor k_1 at the crack tip is given by the result

$$k_1 = \lim_{r \to b^+}[2(b - r)]^{1/2}\sigma_{zz}^{(m)}(r, 0). \tag{72}$$

Omitting details it can be shown that

$$\frac{k_1}{k_0} = -\frac{[H(b) - G(b)]}{b} \tag{73}$$

where k_0 is the stress intensity factor for a penny-shaped crack of radius b which is located in a homogeneous elastic solid

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and subjected to a uniform strain ϵ_0 (or uniform stress $E_m\epsilon_0$) (Sneddon, 1946), i.e.,

$$k_0 = \frac{2\sigma_0\sqrt{b}}{\pi}. \tag{74}$$

Numerical Results

The numerical procedure described in Appendix B is used to determine the stress intensity factor at the tip of the penny-shaped crack which is located in the matrix region. The accuracy of the numerical scheme was verified by comparison with exact analytical results. For example the result for the stress intensity factor for a penny-shaped crack located in an isotropic elastic solid can be recovered in the limit $E_f = E_m = E$; $\nu_f = \nu_m = \nu$. This reduction can also be confirmed from the result obtained from the general solution in the case where $a \to 0$. The numerical result for this particular case compares very accurately, to within 0.01 percent, with the exact closed-form result (74). The material characteristics influencing the stress intensity factor k_1 include, the fiber-matrix modular ratio (E_f/E_m), the Poisson's ratios (ν_f, ν_m) and the fiber-crack radii ratio (a/b). It is of interest to evaluate the amplification of the stress intensity factors at the matrix crack for typical fiber-reinforced composites encountered in engineering applications. Results for several fiber-reinforced composites consisting of E-Glass, Kevlar, Stainless Steel, fibers and matrices consisting of Epoxy Resin, reaction bonded Silicon Nitride (Si_3N_4) and aluminum have been computed using the basic elastic properties of the fiber and matrix planes indicated in Table 1. The elastic properties for these constituents are given by Lawn and Wilshaw (1975), Davidge (1979), Sih and Chen (1981), Hull (1981), and Watt and Perov (1985). Figure 2 indicates the variation of k_1/k_0 for typical fiber-reinforced composites. These results also emphasize the influence of E_f/E_m on the amplification of k_1/k_0.

In all computations, the geometrical parameter a/b is restricted to the range $(a/b)\epsilon(0, 0.90)$. In the limit as $a/b \to 1$ the matrix-fiber crack problem reduces to that of a cracked fiber which is embedded in an undamaged matrix. In this instance, the character of the crack problem changes due to the fact that the crack tip now terminates at an interface (Zak and Williams, 1963). The stress singularity can now exhibit an order different from $1/r^{1/2}$ . . . depending upon the magnitudes of the elastic properties of the fiber and matrix regions. In order to solve the resulting problem, it is of course necessary to reformulate the integral equation scheme where the specific character of the singular stress fields are incorporated in the analysis. This analysis will be examined in detail in a subsequent paper.

Conclusions

Fiber breakage in particular will occur in a fiber-reinforced solid and such discontinuities can introduce complex stress states in the matrix region. In the presence of strong fiber-matrix interface adhesion, the crack extension will occur into the matrix. Although the analysis of crack extension should take into consideration the influence of other intact fibers in the vicinity of the cracked fiber location, some insight into the micromechanical processes of matrix crack extension can be gained by examining a *reduced* problem related to an isolated fiber. Such an idealized isolated fiber configuration is also relevant to the study of matrix crack extension in "fragmentation tests" which are developed for the study of fiber-matrix interface adhesion. The paper has developed a mathematical formulation of a fiber-matrix crack which is subjected to a homogeneous strain field. In this study attention has been restricted to the idealized problem of a penny-shaped matrix crack which is located at a cracked fiber. It is shown that this idealized fiber-matrix crack problem can be formulated within the context of the classical theory of elasticity to evaluate results of interest to microfracture mechanics of composite materials.

The numerical results derived for typical fiber-reinforced composite materials with $E_f > E_m$ indicate that the stress intensity factor at the matrix crack boundary is always higher than that for a penny-shaped crack of equal radius located in a homogeneous matrix (without a fiber). This stress intensity factor amplification increases dramatically with increase in the fiber-matrix modular ratio. The condition for brittle elastic crack extension can be obtained by comparing the stress intensity factor for the fiber-matrix crack with the critical value of the stress intensity factor k_{Ic} applicable to the matrix material. Alternatively, the results given in this study can be used to assess the stable geometries of fiber-matrix cracks that can occur at cracked fiber locations.

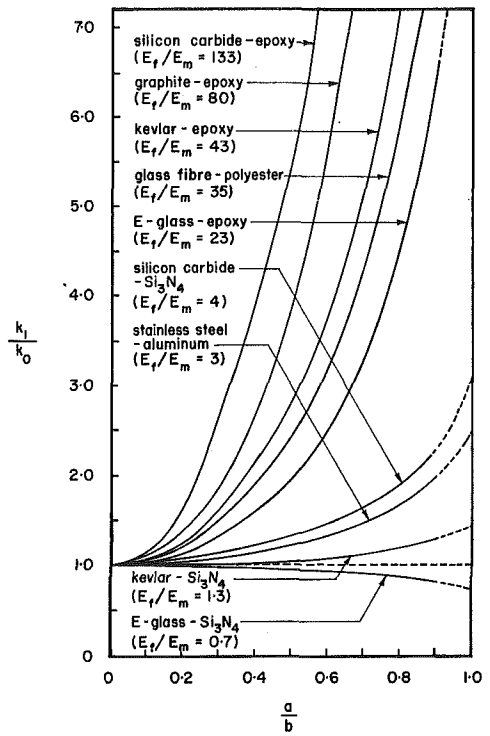


Fig. 2 Stress intensity factor for the fiber-matrix crack. Results for typical fiber-reinforced composites.

Table 1 Elastic constants for components of fiber–reinforced composites

Material	Young's modulus (GN/m^2)	Poisson's ratio
Epoxy resin	3.0	0.40
Si_3N_4	100.0	0.20
Silicon carbide	400.0	0.20
E-glass	70.0	0.26
Kevlar	130.0	0.32
Polyester	2.0	0.40
Glass fibre	70.0	0.20
Graphite	240.0	0.30
Stainless steel	207.0	0.30
Aluminum	69.0	0.33

Acknowledgments

The research described in this paper was supported by a Research Grant A3866 awarded by the Natural Sciences and Engineering Research Council of Canada to the first author.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

Atkinson, C., 1979, "Stress singularities and fracture mechanics," ASME *Applied Mechanics Reviews,* Vol. 32, pp. 123–135.

Atkinson, K., 1976, *A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind,* Soc. Industr. Appl. Math., Philadelphia, PA.

Aveston, J., Cooper, G. A., and Kelly, A., 1971, "Single and multiple fracture," *Conf. Properties of Fibre Composites,* National Physical Laboratory, IPC Science and Technology Press, pp. 15–26.

Backlund, J., 1981, Fracture analysis of notched composites, *Computers and Structures,* Vol. 13, pp. 145–154.

Baker, C. T. H., 1977, *The Numerical Treatment of Integral Equations,* Clarendon Press, Oxford.

Budiansky, B., Hutchinson, J. W., and Evans, A. G., 1986, "Matrix Fracture in Fiber-Reinforced Ceramics," *Journal of the Mechanics and Physics of Solids,* Vol. 34, pp. 167–189.

Busschen, A. ten, and Selvadurai, A. P. S., 1995, "Mechanics of the Segmentation of an Embedded Fiber, Part I: Experimental Investigations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 62, pp. 87–97.

Chamis, C., 1974, "Mechanics of load transfer at the interface," *Composite Materials, Interfaces in Polymer Matrix Composites,* E. P. Plueddemann, ed., Academic Press, New York, pp. 31–77.

Chen, E. P., and Sih, G. C., 1971, "Interfacial delamination of a layered composite under anti-plane strain," *J. Composite Materials,* Vol. 5, pp. 12–23.

Cherapanov, G. P., 1979, *Mechanics of Brittle Fracture,* McGraw-Hill, New York.

Comninou, M., and Schmueser, D., 1979, "The interface crack in a combined tension-compression and shear field," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 345–348.

Davidge, R. W., 1979, *Mechanical Behaviour of Ceramics,* Cambridge University Press, Cambridge, U.K.

Delves, L. M., and Mohamed, J. L., 1985, *Computational Methods for Integral Equations,* Cambridge University Press, Cambridge, U.K.

Dhaliwal, R. S., Singh, B. M., and Rokne, J., 1979, "Penny-shaped crack in an infinite elastic cylinder bonded to an infinite elastic material surrounding it," *Int. J. Eng. Sci.,* Vol. 17, pp. 1245–1255.

Dundurs, J., and Mura, T., 1964, "Interaction between edge dislocation and an circular inclusion," *J. Mech. Phys. Solids,* Vol. 12, pp. 177–189.

Dvorak, G. J., ed., 1991, *Inelastic Deformations of Composite Materials,* Proc. IUTAM Symposium, Troy, NJ, Springer-Verlag, Berlin.

England, A. H., 1965, "A crack between dissimilar media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 400–404.

Erdogan, F., and Ozbeck, T., 1969, "Stresses in fiber-reinforced composites with imperfect bonding," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 865–869.

Hull, D., 1981, *An Introduction to Composite Materials,* Cambridge University Press, Cambridge, U.K.

Karihaloo, B. L., and Viswanathan, K., 1985, "Elastic field of an elliptic inhomogeneity with debonding over an arc (Antiplane strain)," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 91–97.

Kelly, A., 1970, "Interface effects and the work of fracture of a fibrous composite," *Proc. Roy. Soc., London,* Vol. A319, pp. 95–116.

Lawn, B. R., and Wilshaw, T. R., 1975, *Fracture of Brittle Solids,* Cambridge University Press, Cambridge, U.K.

Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity,* Dover, New York.

Marshall, D. B., Cox, B. N., and Evans, A. G., 1985, "The mechanics of matrix cracking in brittle-matrix fibre composites," *Acta Metallurgica,* Vol. 33, pp. 2013–2021.

McCartney, L. N., 1987, "Mechanics of matrix cracking in brittle-matrix fibre reinforced composites," *Proc. Roy. Soc. London,* Vol. A409, pp. 329–350.

Mura, T., 1988, "Inclusion problems," ASME *Applied Mechanics Reviews,* Vol. 41, pp. 15–20.

Mura, T., 1989, *Micromechanics of Defects in Solids,* 2nd ed., Sijthoff and Noordhoff, The Netherlands.

Narayan, T. V., and Erdogan, F., 1975, "Penny-shaped crack in a fiber reinforced matrix," *Int. J. Solids Struct.,* Vol. 11, pp. 1315–1327.

Ozbek, T., and Erdogan, F., 1969, "Some elasticity problems in fibre-reinforced composites with imperfect bonds," *Int. J. Engng. Sci.,* Vol. 7, pp. 931–946.

Pacella, A. H., and Erdogan, F., 1974, "A penny-shaped crack in a filament reinforced matrix: II. The crack problem," *Int. J. Engng. Sci.,* Vol. 10, pp. 807–819.

Rose, L. R. F., 1987, "Crack reinforcement by distributed springs," *J. Mech. Phys. Solids,* Vol. 35, pp. 383–405.

Selvadurai, A. P. S., 1981, "A theoretical model for elastic flaw bridging in unidirectional fibre reinforced composites," *Mechanics of structured Media, Studies in Applied Mechanics,* A. P. S. Selvadurai, ed., Vol. 5A, pp. 59–72.

Selvadurai, A. P. S., 1991a, "Cracks at the extremities of cylindrical fibre inclusion," *Proc. IUTAM Symp. Inelastic Deformation of Composite Materials,* G. J. Dvorak, ed., Troy, New York, Springer-Verlag, Berlin, pp. 147–171.

Selvadurai, A. P. S., Singh, B. M., and Au, M. C., 1991, "The in-plane loading of a rigid disk inclusion embedded in an elastic halfspace," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 362–369.

Selvadurai, A. P. S., and Busschen, A. ten, 1995, "Mechanics of the Segmentation of an Embedded Fiber, Part II: Computational Modeling and Comparisons," ASME JOURNAL OF APPLIED MECHANICS, Vol. 62, pp. 98–107.

Sendeckyi, G. P., 1974, "Debonding of rigid curvilinear inclusions in longitudinal shear deformations," *Engineering Fracture Mechanics,* Vol. 6, pp. 33–45.

Sendeckyi, G. P., 1975, "Some topics in anisotropic elasticity," *Composite Materials, Structural Design and Analysis,* Part I, C. C. Chamis, ed., Academic Press, New York, pp. 1–50.

Sih, G. C., 1979, "Fracture mechanics of composite materials," *Proc. 1st USA-USSR Symposium on Fracture of Composite Materials,* G. C. Sih and V. P. Tamuzs, eds., Sijthoff and Noordhoff, The Netherlands, pp. 113–130.

Sih, G. C., and Chen, E. P., 1981, *Cracks in Composite Materials, Mechanics of Fracture,* Vol. 6, Martinus Nijhoff, The Hague.

Singh, B. M., Dhaliwal, R. S., and Rokne, J., 1979, "An axisymmetric external crack problem for an infinite medium with a cylindrical inclusion," *J. Engng. Math.,* Vol. 13, pp. 235–247.

Sneddon, I. N., 1946, "The distribution of stress in the neighbourhood of a crack in an elastic solid," *Proc. Roy. Soc., London,* Vol. A187, pp. 229–260.

Sneddon, I. N., 1951, *Fourier Transforms,* McGraw-Hill, New York.

Stang, H., 1987, "A double-inclusion model for microcrack arrest in fibre reinforced brittle materials," *J. Mech. Phys. Solids,* Vol. 35, pp. 325–343.

Taya, M., and Mura, T., 1981, "On stiffness and strength of an aligned containing fiber-end cracks under uniaxial applied stress," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 361–367.

Taya, M., and Chou, T. W., 1983, "Energy release rates of various microcracks in short-fibre composites," *Mechanics of Composite Materials, Proc. IUTAM Symp.,* C. T. Herakovich and Z. Hashin, eds., Pergamon Press, pp. 457–472.

Timoshenko, S., and Goodier, J. N., 1970, *Theory of Elasticity,* 3rd ed., McGraw-Hill, New York.

van den Berg, R. M., 1990, *Fragmentatleproeven net Elastvezel/Polyester en Glasvezel/Epoxy,* Report LTM-918, Tech. Univ., Delft.

Watt, W., and Perov, B. V., 1985, *Strong Fibres, Handbook of Composites,* Vol. I, A. Kelly and Yu. N. Rabotnow, eds., North Holland, Amsterdam.

Wijeywickrema, A. C., Keer, L. M., Harashima, K., and Mura, T., 1990, "The annular crack surrounding an elastic fibre in a tension field," *Int. J. Solids Struct.,* Vol. 21, pp. 315–328.

Zak, A. R., and Williams, M. L., 1963, "Crack Point Stress Singularities at Bimaterial Interface," ASME JOURNAL OF APPLIED MECHANICS, Vol. 30, pp. 142–143.

APPENDIX A

The expressions for the functions $A(s)$, $B(s)$, $C(s)$, and $D(s)$, take the following forms:

$$\zeta A(s) = -\Big[X_1(s)\{(a_{13}I_0 + a_8I_1)(a_1a_6 - a_3a_4) + a_3I_1(a_{12}I_0 + a_7I_1)\} + X_2(s)\{(a_9I_1 + a_{14}I_0)(a_1a_6 - a_3a_4) + a_3I_0(a_{12}I_0 + a_7I_1)\} + X_3(s)\{(a_{10}I_1 + a_{15}I_0)(a_1a_6 - a_3a_4) - a_6I_1(a_{12}I_0 + a_7I_1)\} + X_4(s)\Big\{(a_{11}I_1 + a_{16}I_0) \times (a_1a_6 - a_3a_4) + \frac{a_6}{s}(I_1 - asI_0)(a_{12}I_0 + a_7I_1)\Big\}\Big] \quad \text{(A1)}$$

$$\zeta B(s) = X_1(s)\{a_2I_1(a_{12}I_0 + a_7I_1) - (a_2a_4 - a_1a_5) \times (a_{13}I_0 + a_8I_1)\} + X_2(s)\{a_2I_0(a_{12}I_0 + a_7I_1) - (a_2a_4 - a_1a_5)(a_{14}I_0 + a_9I_1)\} - X_3(s) \times \{a_5I_1(a_{12}I_0 + a_7I_1) + (a_2a_4 - a_1a_5)(a_{15}I_0 + a_{10}I_1)\} + X_4(s)\Big\{\frac{a_5}{s}(I_1 - saI_0)(a_{12}I_0 + a_7I_1) - (a_2a_4 - a_1a_5)(a_{16}I_0 + a_{11}I_1)\Big\} \quad \text{(A2)}$$

$$\eta C(s) = X_1(s)\{a_7a_{13} - a_8a_{12}\} + X_2(s)\{a_7a_{14} - a_{12}a_9\} + X_3(s)\{a_{15}a_7 - a_{10}a_{12}) + X_4(s)\{a_7a_{16} - a_{11}a_{12}\} \quad \text{(A3)}$$

$$\eta D(s) = X_1(s)\{a_{13}I_0 + a_8I_1\} + X_2(s)\{a_9I_1 + a_{14}I_0\} + X_3(s)\{a_{10}I_1 + a_{15}I_0\} + X_4(s)\{a_{11}I_1 + a_{16}I_0\} \quad \text{(A4)}$$

where

$$\zeta = (a_2a_6 - a_3a_5)(a_{12}I_0 + a_7I_1) \quad \text{(A5)}$$

$$\eta = (a_7I_1 + a_{12}I_0) \quad \text{(A6)}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$a_1 = s^2a^2(I_1^2 - I_0^2) + 2(1 - \nu_f)I_1^2 \quad \text{(A7)}$$

$$a_2 = sa\Gamma\{K_0I_1 + K_1I_0\} \quad \text{(A8)}$$

$$a_3 = -\Gamma[K_1I_1\{2 - 2\nu_m + s^2a^2\} + 2(1 - \nu_m)saI_0K_1 + 2sa(1 - \nu_m)K_0I_1 + K_0I_0s^2a^2] \quad \text{(A9)}$$

$$a_4 = sa(I_1^2 - I_0^2) + 4(1 - \nu_f)I_0I_1 \quad \text{(A10)}$$

$$a_5 = \frac{a_2}{sa\Gamma} \quad \text{(A11)}$$

$$a_6 = -\{sa(I_1K_1 + I_0K_0) + 4(1 - \nu_m)I_0K_1\} \quad \text{(A12)}$$

$$a_7 = saI_1 + \left\{\frac{saK_1(a_2a_4 - a_5a_1) + K_0(a_3a_4 - a_6a_1)}{a_2a_6 - a_3a_5}\right\} \quad \text{(A13)}$$

$$a_8 = 1 + \frac{I_1(a_2saK_1 + a_3K_0)}{(a_2a_6 - a_3a_5)} \quad \text{(A14)}$$

$$a_9 = \frac{I_0(saa_2K_1 + a_3K_0)}{(a_2a_6 - a_3a_5)} \quad \text{(A15)}$$

$$a_{10} = -\frac{I_1(a_6K_0 + saa_5K_1)}{(a_2a_6 - a_3a_5)} \quad \text{(A16)}$$

$$a_{11} = \left\{\frac{a_6K_0 + saa_5K_1}{s}\right\}\left\{\frac{I_1 - saI_0}{a_2a_6 - a_3a_5}\right\} \quad \text{(A17)}$$

$$a_{12} = 2(1 - \nu_f)I_1 - saI_0 + \frac{\Gamma K_1(a_3a_4 - a_1a_6)}{(a_2a_6 - a_3a_5)} + \frac{\Gamma(a_2a_4 - a_1a_5)}{(a_2a_6 - a_3a_5)}\{2(1 - \nu_m)K_1 + saK_0\} \quad \text{(A18)}$$

$$a_{13} = \frac{\Gamma}{(a_2a_6 - a_3a_5)}[a_3K_1I_1 + a_2I_1\{2(1 - \nu_m)K_1 + saK_0\}] \quad \text{(A19)}$$

$$a_{14} = \frac{\Gamma}{(a_2a_6 - a_3a_5)}[a_3K_1I_1 + a_2I_0\{2(1 - \nu_m)K_1 + saK_0\}] \quad \text{(A20)}$$

$$a_{15} = \frac{-\Gamma I_1}{(a_2a_6 - a_3a_5)}[a_6K_1 + a_5\{2(1 - \nu_m)K_1 + saK_0\}] \quad \text{(A21)}$$

$$a_{16} = \frac{1}{s}\left[1 + \frac{\Gamma(I_1 - saI_0)}{(a_2a_6 - a_3a_5)} \times \{a_6K_1 + [2(1 - \nu_m)K_1 + saK_0]a_5\}\right] \quad \text{(A22)}$$

and

$$\Gamma = \frac{\mu_m}{\mu_f}. \quad \text{(A23)}$$

By making use of the expressions (53) to (60) and Eqs. (32) to (35) we can obtain expressions for the functions $X_i(s) = (i = 1, 2, 3, 4)$ in terms of $h(s)$, $g(s)$, and $m(s)$. These expressions are

$$X_1(s) = [(3 - 2\nu_m)I_0(sa) + saI_1(sa)]\left[\int_a^b h(u)e^{-su}du + \int_b^\infty g(u)e^{-su}du\right] - sI_0(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right] - \frac{2}{\pi}\int_0^a [(3 - 2\nu_f)\sinh(su)K_0(sa) - saK_1(sa)\sinh(su) + su\cosh(su)K_0(sa)]m(u)du \quad \text{(A24)}$$

$$X_2(s) = -\left\{\frac{sa}{2}[I_0(sa) + I_2(sa)] + 2\nu_mI_1(sa)\right\} \times \left\{\int_a^b h(u)e^{-su}du + \int_b^\infty g(u)e^{-su}du\right\} + sI_1(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right] - \frac{2}{\pi}\int_0^a \{(2\nu_f - 1)K_1(sa)\sinh(su) - [sa\sinh(su)K_0(sa) - su\cosh(su)K_1(sa)]\}m(u)du \quad \text{(A25)}$$

$$X_3(s) = \frac{2}{\pi}\left[\int_0^a m(u)\{sa[sa\sinh(su)K_1(sa) - su\cosh(su)K_0(sa)] + (1 - 2\nu_f)K_1(sa)\sinh(su) - suK_1(sa)\cosh(su)\}du\right] + \Gamma\int_a^b h(t)e^{-st} \times \left\{sa\left[\frac{1}{2}I_0(sa) + saI_1(sa) - stI_0(sa)\right] + I_1(sa)(st - 2\nu_m) - \frac{sa}{2}I_2(sa)\right\}dt + \Gamma\int_b^\infty g(t)e^{-st}\left[sa\left\{\frac{1}{2}I_0(sa) + saI_1(sa) - stI_0(sa)\right\} + I_1(sa)\{st - 2\nu_m\} - \frac{sa}{2}I_2(sa)\right]dt \quad \text{(A26)}$$

$$X_4(s) = -\Gamma\left[2sI_1(sa) + \frac{s^2a}{2}\{I_0(sa) + I_2(sa)\}\right] \times \left[\int_a^b h(u)e^{-su}du + \int_b^\infty g(u)e^{-su}du\right] + \Gamma s^2I_1(sa)\left[\int_a^b uh(u)e^{-su}du + \int_b^\infty ug(u)e^{-su}du\right] - \frac{2s}{\pi}\int_0^a m(u)[\sinh(su)K_1(sa) + su\cosh(su)K_1(sa) - sa\sinh(su)K_0(sa)]du. \quad \text{(A27)}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The expressions for $P_i(s, t)$ $(i = 1, 2, 3, 4)$ occurring in the kernel functions S_1 and S_2 defined by Eqs. (65) and (66) are given by

$$\zeta P_1(s,t) = -[\{(a_{12}I_0 + a_7I_1)a_3I_1 + (a_8I_1 + a_{13}I_0) \times (a_1a_6 - a_3a_4)\}L_1(s,t) - \{a_2I_1(a_{12}I_0 + a_7I_1) - (a_2a_4 - a_1a_5)(a_{13}I_0 + a_8I_1)\} \times \{2\nu_m L_1(s,t) - sL_2(s,t)\}] \quad \text{(A28)}$$

$$\zeta P_2(s,t) = -[\{(a_{12}I_0 + a_7I_1)a_3I_0 + (a_9I_1 + a_{14}I_0) \times (a_1a_6 - a_3a_4)\}L_1(s,t) - \{a_2I_0(a_{12}I_0 + a_7I_1) - (a_2a_4 - a_1a_5)(a_{14}I_0 + a_9I_1)\} \times \{2\nu_m L_1(s,t) - sL_2(s,t)\}] \quad \text{(A29)}$$

$$\zeta P_3(s,t) = [\{(a_{12}I_0 + a_7I_1)a_6I_1 - (a_{10}I_1 + a_{15}I_0) \times (a_1a_6 - a_3a_4)\}L_1(s,t) - \{a_5I_1(a_{12}I_0 + a_7I_1) + (a_2a_4 - a_1a_5)(a_{15}I_0 + a_{10}I_1)\} \times \{2\nu_m L_1(s,t) - sL_2(s,t)\}] \quad \text{(A30)}$$

$$\zeta P_4(s,t) = -\left[\left\{\frac{a_6}{s}(a_{12}I_0 + a_7I_1)(I_1 - saI_0) + (a_1a_6 - a_3a_4)(a_{11}I_1 + a_{16}I_0)\right\}L_1(s,t) - \left\{\frac{a_5}{s}(a_{12}I_0 + a_7I_1)(I_1 - saI_0) - (a_2a_4 - a_1a_5) \times (a_{16}I_0 + a_{11}I_1)\right\} \cdot \{2\nu_m L_1(s,t) - sL_2(s,t)\}\right]. \quad \text{(A31)}$$

The expressions for $R_i(s, t)$ $(i = 1, 2, 3, 4)$ occurring in the kernel functions S_3 and S_4 defined by Eqs. (68) and (69) are given by

$$\eta R_1(s,t) = [\sinh(st)\{a_7a_{13} - a_8a_{12}\} - (a_{13}I_0 + a_8I_1) \times \{\sinh(st)(2\nu_f - 1) + st\cosh(st)\}] \quad \text{(A32)}$$

$$\eta R_2(s,t) = [\sinh(st)\{a_7a_{14} - a_9a_{12}\} - (a_{14}I_0 + a_9I_1) \times \{\sinh(st)(2\nu_f - 1) + st\cosh(st)\}] \quad \text{(A33)}$$

$$\eta R_3(s,t) = [\sinh(st)\{a_7a_{15} - a_{10}a_{12}\} - (a_{15}I_0 + a_{10}I_1) \times \{\sinh(st)(2\nu_f - 1) + st\cosh(st)\}] \quad \text{(A34)}$$

$$\eta R_4(s,t) = [\sinh(st)\{a_7a_{16} - a_{11}a_{12}\} - (a_{16}I_0 + a_{11}I_1) \times \{\sinh(st)(2\nu_f - 1) + st\cosh(st)\}]. \quad \text{(A35)}$$

APPENDIX B

The system of integral Eqs. (61), (67), and (70) can be represented in the generalized form

$$\mathbf{X}(t) + \int_0^\infty \mathbf{X}(u)\mathbf{K}(u,t)du = \mathbf{B}(t); \quad 0 \le t < \infty \quad \text{(B1)}$$

where the unknown functions $X(t)$ are defined as

$$\mathbf{X}(t) = \begin{cases} M(t); & 0 \le t < a \\ H(t); & a \le t < b \\ G(t); & b \le t < \infty. \end{cases} \quad \text{(B2)}$$

The right-hand side of the prescribed function $\mathbf{B}(t)$ in (B1) takes the form

$$\mathbf{B}(t) = \begin{cases} t\Gamma\Omega; & 0 \le t < a \\ (b^2 - t^2)^{1/2}; & a \le t < b \\ 0; & b \le t < \infty \end{cases} \quad \text{(B3)}$$

and the kernel functions $\mathbf{K}(u, t)$ are defined as follows:

$$\mathbf{K}(u,t) = \begin{cases} \frac{4}{\pi^2}S_4(u,t); & (0 < u < a; 0 < t < a) \\ \frac{2}{\pi}S_3(u,t); & (a < u < \infty; 0 < t < a) \\ \frac{4}{\pi^2}S_2(u,t); & (0 < u < a; a < t < b) \\ \frac{2}{\pi}S_1(u,t); & (a < u < b; a < t < b) \\ \frac{2}{\pi}S_1(u,t) & \\ -\frac{2}{\pi}\frac{(b^2 - t^2)^{1/2}}{(u^2 - t^2)(u^2 - b^2)^{1/2}}; & (b < u < \infty; a < t < b) \\ -\frac{2}{\pi}\frac{t(b^2 - u^2)^{1/2}}{(t^2 - b^2)^{1/2}(t^2 - u^2)}; & (a < u < b; b < t < \infty), \end{cases} \quad \text{(B4)}$$

with $\mathbf{K}(u, t)$ being zero for all other intervals of $u\epsilon(0, \infty)$, $t\epsilon(0, \infty)$.

The system of integral Eqs. (B1) are systems of coupled Fredholm integral equations of the second kind. The solution of this class of integral equations via numerical techniques has been extensively documented in the texts by Atkinson (1976), Baker (1977), and Delves and Mohamed (1985). Examples of applications of these procedures are also given by Selvadurai et al. (1991).

For the numerical solution of the Eq. (B1) we discretize the intervals $[0, a]$, $[a, b]$ and $[b, \infty]$ by N_1, N_2, and N_3 segments, respectively. The corresponding end points of the segments are given by

$$x_i = (i - 1)h_x; \quad i = 1, 2, \ldots (N_1 + 1) \quad \text{(B5)}$$

$$y_i = a + (i - 1)h_y; \quad i = 1, 2, \ldots (N_2 + 1) \quad \text{(B6)}$$

$$z_i = z_{i-1} + C_r(z_{i-1} - z_{i-2}); \quad i = 3, 4, \ldots (N_3 + 1) \quad \text{(B7)}$$

with $z_1 = b$ and $z_2 = b + h_y$ where h_x and h_y are given by a/N_1 and $(b - a)/N_2$, respectively, and C_r is a constant greater than unity such that the recurrence relation (B7) can generate a large number at z_{N_3+1}. For the collocation points of the integral equations we have

$$t_i = \begin{cases} (x_i + x_{i+1})/2; & 1 \le i \le N_1 \quad \text{(B8)} \\ (y_i + y_{j+1})/2; & N_1 < i < N_1 + N_2; \\ & j = i - N_1 \quad \text{(B9)} \\ (z_j + z_{j+1})/2; & N_1 + N_2 < i < N; \\ & j = i - N_1 - N_2 \quad \text{(B10)} \end{cases}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where $N = N_1 + N_2 + N_3$. Using the above procedures and representations, the discretized form of (B1) can be written as

$$\mathbf{A}_{ij}\mathbf{X}_j = \mathbf{B}_j \tag{B11}$$

where the unknowns are

$$\mathbf{X}_i = \mathbf{X}(t_i). \tag{B12}$$

The right-hand side of the vector is

$$\mathbf{B}_i = \mathbf{B}(t_i) \tag{B13}$$

and the coefficients of the matrix $\mathbf{A}_{ij}$ are

$$\mathbf{A}_{ij} = \delta_{ij} + K(t_i, t_j)\Delta t, \tag{B14}$$

where δ_{ij} is the Kronecker delta and Δt is given as h_x, h_y or $(z_{j+1} - z_j)$ with $1 < j \leq N_1$, $N_1 < j \leq (N_1 + N_2)$ or $(N_1 + N_2) < j \leq N$, respectively.

Upon solution of the matrix Eq. (B11), the stress intensity factor at the tip of the crack in the matrix region can be obtained from (73); i.e.,

$$\frac{k_1}{k_0} = \frac{\{X_{N_1+N_2+1} - X_{N_1+N_2}\}}{b}. \tag{B15}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Consistent Infinitesimal Finite Element Cell Method: Three-Dimensional Scalar Wave Equation

J. P. Wolf

Chongmin Song

Institute of Hydraulics and Energy,
Department of Civil Engineering,
Swiss Federal Institute of Technology,
Lausanne CH-1015, Switzerland

To calculate the unit-impulse response matrix of an unbounded medium governed by the scalar wave equation for use in a time-domain analysis of unbounded medium-structure interaction, the consistent infinitesimal finite element cell method is developed for the three-dimensional case. Its derivation is based on the finite element formulation and on similarity. The discretization is only performed on the structure-medium interface, yielding a reduction of the spatial dimension by 1. The procedure is exact in the radial direction and converges to the exact solution in the finite element sense in the circumferential directions.

Introduction

To analyze unbounded medium-structure interaction such as occurring in linear acoustics and fluid dynamics, an accurate representation of the unbounded medium is required. In the substructure method of analysis in the time domain (where the structure and the unbounded medium are first addressed separately), the unit-impulse response matrix corresponding to the discretization on the structure-medium interface of the unbounded medium is determined before a transient can be processed. As an alternative to the boundary element method (Friedman and Shaw, 1962; Dominguez, 1993), which applies an analytical solution incorporating the condition at infinity, the *consistent infinitesimal finite element cell method* based solely on the finite element formulation has been developed for the scalar wave equation in two dimensions (Song and Wolf, 1995). Explicit expressions for the coefficient matrices of a two-node line finite element on the interface are specified. The goal of this paper is to expand the formulation to the three-dimensional scalar wave equation and to introduce any type of surface finite elements for the discretization of the structure-medium interface. In this case the coefficient matrices are calculated by numerical integration on the interface.

The fundamental idea of the consistent infinitesimal finite element cell method is illustrated. In Fig. 1 the unbounded (semi-infinite) medium taking the irregular structure-medium interface into account is shown. The unbounded medium with the characteristic length r_i is to be analyzed. To be able to address the unbounded medium with another *similar* structure-medium interface with the characteristic length r_e, a cell consisting of finite elements is introduced. Its interior boundary coincides with the analyzed structure-medium interface (subscript i for interior boundary) and its similar exterior boundary is constructed from the interior boundary by applying the similarity factor r_e/r_i (subscript e for exterior boundary). The exterior boundary of the cell coincides with the interface of the unbounded medium with r_e. Adding the bounded cell of finite elements to the unbounded medium with r_e results in the similar unbounded medium with r_i. This concept can be applied to their dynamic-stiffness matrices in the frequency domain (Dasgupta, 1982). Assembling the dynamic-stiffness matrix of the cell (which is straightforwardly determined from its static-stiffness and mass matrices) and the unknown dynamic-stiffness matrix of the unbounded medium characterized by the length r_e results in the unknown dynamic-stiffness matrix of the unbounded medium with length r_i. As a relationship for the dynamic-stiffness matrices defined at similar interfaces characterized by different lengths of the unbounded medium exists, the procedure leads to an expression for the dynamic-stiffness matrix of the unbounded medium as a function of the property matrices of the cell. Analogously, this applies also to the unit-impulse response matrix. The cell width measured in the radial direction is selected as infinitesimal, whereby the limit is performed analytically. This results in the consistent infinitesimal finite element cell method, which is a stand-alone finite element formulation capable of capturing the radiation condition at infinity without using analytical solutions. The formulation will depend only on the geometry of the structure-medium interface and on the material properties of the unbounded medium, yielding a reduction of the spatial dimension by one. As the limit of the infinitesimal cell width in the radial direction is performed analytically, the error of discretization in this direction vanishes. In the circumferential directions the error of the finite element discretization remains. As the lengths of the finite elements become infinitesimal, this error vanishes. Thus, this procedure is *exact* in the radial direction and *converges to the exact solution in the finite element sense* in the circumferential directions. For problems with a boundary extending from the structure-medium interface to infinity (such as a half-space with a free surface), this novel method automatically incorporates this boundary condition during the assemblage process in contrast to the boundary element method. Material inhomogeneities which satisfy similarity can also be processed without any additional effort. Only the coefficient matrices which follow from the analytical limit of the conventional static-stiffness and mass matrices of the bounded finite-element cell need to be calculated, and they are then used in standard matrix operations to obtain the unit-impulse response matrix of the unbounded medium.

The well-known consistent boundary method (e.g., Kausel, Roësset, and Waas, 1975) to analyze a two-dimensional horizontally layered medium (in the frequency domain) is a special case of the consistent infinitesimal finite element cell method (in the time domain).

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 13, 1995; final revision, Mar. 7, 1996. Associate Technical Editor: W. L. Liu.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

To avoid unnecessary duplication with Song and Wolf (1995) only the coefficient matrices of any type of finite element in the three-dimensional case are derived. All other aspects such as the derivation of the consistent infinitesimal finite element cell equation and its numerical implementation are covered in detail in Song and Wolf (1995) which the reader should consult.

Finite Element Discretization

The scalar wave equation in three dimensions is formulated as

$$u_{,xx} + u_{,yy} + u_{,zz} - \frac{\ddot{u}}{c^2} = 0. \tag{1}$$

$u = u(x, y, z, t)$ denotes the function to be determined, and c is the wave velocity. Besides vanishing initial conditions, boundary conditions are prescribed as

$$u = \bar{u} \quad \text{on} \quad S_u \tag{2a}$$

$$u_{,n} = \bar{v} \quad \text{on} \quad S_v. \tag{2b}$$

$\bar{u}$ is the prescribed function on the surface S_u and $\bar{v}$ the prescribed normal flux on S_v.

For acoustic waves u corresponds to the velocity potential, and c is the speed of sound in the acoustic medium. The velocity vector and the pressure follow as $\{v\} = [u_{,x}\, u_{,y}\, u_{,z}]^T$ and $p = -\rho\dot{u}$, respectively (mass density ρ). In Eq. (2b) $\bar{v}$ corresponds to the velocity in the normal direction.

After the spatial discretization of the finite element method the semidiscrete wave equation of the global system is formulated as

$$[M]\{\ddot{u}(t)\} + [K]\{u(t)\} = \{P(t)\}. \tag{3}$$

$\{u(t)\}$ denotes the vector of the nodal values of the function. $[K]$, $[M]$, and $\{P(t)\}$ are the global static-stiffness matrix, mass matrix, and the nodal flux vector, which are determined by assembling the corresponding quantities on an element level. On the structure-medium interface of an unbounded medium the nodal flux vector can be written as the convolution integral with the second time derivative of the nodal values of the function

$$\{R(t)\} = \int_0^t [M^\infty(t - \tau)]\{\ddot{u}(\tau)\}d\tau. \tag{4}$$

$[M^\infty(t)]$ is the response matrix to a unit impulse of $\{\ddot{u}(t)\}$. Equation (4) corresponds to Eq. (58) of Song and Wolf (1995).

Coefficient Matrices of Finite Element Cell

In a three-dimensional unbounded medium-structure-interaction analysis the structure-medium interface is, in general, a doubly curved finite surface, which is discretized using a two-dimensional mesh of isoparametric surface finite elements

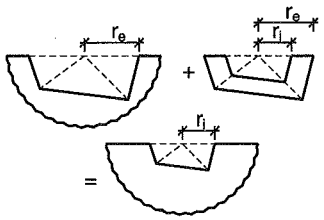


Fig. 1 Fundamental concept of consistent infinitesimal finite element cell method

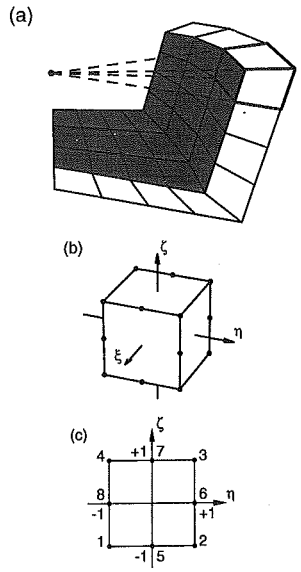


Fig. 2 (*a*) Three-dimensional finite element cell, (*b*) three-dimensional parent element, (*c*) two-dimensional parent element on discretized interface

(shaded area in Fig. 2(a)). Starting from this mesh a three-dimensional finite element cell satisfying similarity with a single element in the radial direction is constructed for the derivation of the consistent infinitesimal finite element cell equation. In the radial direction of the finite element of the cell a linear shape function is introduced. To construct the static-stiffness and mass matrices of the isoparametric three-dimensional finite element, its parent element is addressed. As an example the parent element with eight nodes on the structure-medium interface is shown in Fig. 2(b). The ξ-axis of the parent element corresponds to the radial direction which points from the structure-medium interface towards infinity. At the interior boundary (which coincides with the structure-medium interface) and at the exterior boundary the values $\xi_i = -1$ and $\xi_e = +1$, respectively, result. On the interior boundary $\{N(\eta, \zeta)\}$ denotes the shape functions of the two-dimensional surface finite element. The corresponding eight-node quadrilateral parent element is shown in Fig. 2(c).

The shape functions of the three-dimensional parent element $\{\hat{N}(\xi, \eta, \zeta)\} = \{\hat{N}\}$ (the symbol circumflex $\hat{}$ denotes the three-dimensional finite element) are generated from those of the two-dimensional parent element $\{N(\eta, \zeta)\} = \{N\}$. After decomposition with respect to the interior and exterior boundaries

$$\{\hat{N}\} = \begin{Bmatrix} \{\hat{N}\}_i \\ \{\hat{N}\}_e \end{Bmatrix} \tag{5}$$

applies which are equal to

$$\{\hat{N}\}_j = \tfrac{1}{2}(1 + \xi_j\xi)\{N\} \quad (j = i, e) \tag{6}$$

with a typical element

$$\hat{N}_{jk} = \tfrac{1}{2}(1 + \xi_j\xi)N_k \quad (j = i, e; k = 1, 2, \ldots). \tag{7}$$

j denotes the interior or exterior boundary and k the node number of the two-dimensional parent element. Due to similarity, the coordinates of the nodes on the exterior boundary $\{x\}_e$, $\{y\}_e$, $\{z\}_e$ can be expressed by those of the nodes on the interior boundary $\{x\}$, $\{y\}$, $\{z\}$ and the dimensionless cell width w as

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\{x\}_e = (1 + w)\{x\} \tag{8a}$$

$$\{y\}_e = (1 + w)\{y\} \tag{8b}$$

$$\{z\}_e = (1 + w)\{z\}. \tag{8c}$$

Note that the subscript i for the coordinates of the nodes on the interior boundary, the structure-medium interface, is omitted (as in the actual application only the structure-medium interface is discretized). The nodal coordinates of the three-dimensional finite element $\{\hat{x}\}$, $\{\hat{y}\}$, $\{\hat{z}\}$ are constructed by assembling those of the nodes on the interior and exterior boundaries $\{x\}$ and $\{x\}_e$, $\{y\}$ and $\{y\}_e$, $\{z\}$ and $\{z\}_e$.

Introducing the isoparametric mapping rule and with Eqs. (5), (6), and (8)

$$\hat{x} = \{\hat{N}\}^T\{\hat{x}\} = \{\hat{N}\}_i^T\{x\} + \{\hat{N}\}_e^T\{x\}_e$$

$$= \left(1 + \frac{w}{2}(1 + \xi)\right)\{N\}^T\{x\} \tag{9a}$$

$$\hat{y} = \left(1 + \frac{w}{2}(1 + \xi)\right)\{N\}^T\{y\} \tag{9b}$$

$$\hat{z} = \left(1 + \frac{w}{2}(1 + \xi)\right)\{N\}^T\{z\} \tag{9c}$$

follow. The Jacobian matrix equals

$$[\hat{J}] = \begin{bmatrix} \hat{x}_{,\xi} & \hat{y}_{,\xi} & \hat{z}_{,\xi} \\ \hat{x}_{,\eta} & \hat{y}_{,\eta} & \hat{z}_{,\eta} \\ \hat{x}_{,\zeta} & \hat{y}_{,\zeta} & \hat{z}_{,\zeta} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{w}{2} & & \\ & 1 + \frac{w}{2}(1 + \xi) & \\ & & 1 + \frac{w}{2}(1 + \xi) \end{bmatrix} [J] \tag{10}$$

where the abbreviation

$$[J] = \begin{bmatrix} \{N\}^T\{x\} & \{N\}^T\{y\} & \{N\}^T\{z\} \\ \{N_{,\eta}\}^T\{x\} & \{N_{,\eta}\}^T\{y\} & \{N_{,\eta}\}^T\{z\} \\ \{N_{,\zeta}\}^T\{x\} & \{N_{,\zeta}\}^T\{y\} & \{N_{,\zeta}\}^T\{z\} \end{bmatrix} \tag{11}$$

is introduced. The determinant of the Jacobian matrix equals

$$|\hat{J}| = \frac{w}{2}\left(1 + \frac{w}{2}(1 + \xi)\right)^2 |J|. \tag{12}$$

The gradients of the shape functions equal

$$\{B\}_{jk} = \begin{Bmatrix} \hat{N}_{jk,\hat{x}} \\ \hat{N}_{jk,\hat{y}} \\ \hat{N}_{jk,\hat{z}} \end{Bmatrix} = [\hat{J}]^{-1} \begin{Bmatrix} \hat{N}_{jk,\xi} \\ \hat{N}_{jk,\eta} \\ \hat{N}_{jk,\zeta} \end{Bmatrix}$$

$$= \frac{\xi_j}{w}\{B_p\}_k + \frac{1 + \xi_j\xi}{2\left(1 + \frac{w}{2}(1 + \xi)\right)}\{B_q\}_k \tag{13}$$

where

$$\{B_p\}_k = [J]^{-1} \begin{Bmatrix} N_k \\ 0 \\ 0 \end{Bmatrix}; \quad \{B_q\}_k = [J]^{-1} \begin{Bmatrix} 0 \\ N_{k,\eta} \\ N_{k,\zeta} \end{Bmatrix}. \tag{14}$$

Assembling $\{B_p\}_k$ and $\{B_q\}_k$ of all nodes yields

$$[B_p] = [J]^{-1} \begin{bmatrix} \{N\}^T \\ 0 \\ 0 \end{bmatrix}; \quad [B_q] = [J]^{-1} \begin{bmatrix} 0 \\ \{N_{,\eta}\}^T \\ \{N_{,\zeta}\}^T \end{bmatrix}. \tag{15}$$

$[B_p]$ and $[B_q]$ are not a function of ξ. $[B]$ is decomposed with respect to the interior and exterior boundaries as

$$[B] = [[B]_i[B]_e] \tag{16}$$

with

$$[B]_j = \frac{\xi_j}{w}[B_p] + \frac{1 + \xi_j\xi}{2\left(1 + \frac{w}{2}(1 + \xi)\right)}[B_q] \quad (j = i, e). \tag{17}$$

The static-stiffness matrix of the three-dimensional finite element equals

$$[K] = \int_V [B]^T[B]dV, \tag{18}$$

with the integration over the volume V of the element. $[K]$ is decomposed into submatrices with respect to the interior and exterior boundaries as ($j = i, e$; $l = i, e$)

$$[K]_{jl} = \int_V [B]_j^T[B]_l dV$$

$$= \int_{-1}^{+1}\int_{-1}^{+1}\int_{-1}^{+1} [B]_j^T[B]_l|\hat{J}|\,d\xi d\eta d\zeta. \tag{19}$$

The integration in the ξ-direction is performed analytically. Decomposition with respect to the dimensionless cell width w for later use in the derivation of the consistent infinitesimal finite element cell equation yields

$$[K]_{jl} = \frac{1}{w}[K_0]_{jl} + [K_1]_{jl} + w[K_2]_{jl} \tag{20}$$

where

$$[K_0]_{jl} = \xi_j\xi_l[F] \tag{21a}$$

$$[K_1]_{jl} = \xi_j\xi_l[F] + \frac{\xi_l}{2}[B_1] + \frac{\xi_j}{2}[B_1]^T \tag{21b}$$

$$[K_2]_{jl} = \frac{\xi_j\xi_l}{3}[F] + \left(\frac{\xi_l}{4} + \frac{\xi_j\xi_l}{12}\right)[B_1]$$

$$+ \left(\frac{\xi_j}{4} + \frac{\xi_j\xi_l}{12}\right)[B_1]^T + \left(\frac{1}{4} + \frac{\xi_j\xi_l}{12}\right)[K_2]_{sum} \tag{21c}$$

with

$$[F] = \int_{-1}^{+1}\int_{-1}^{+1} [B_p]^T[B_p]|J|\,d\eta d\zeta \tag{22a}$$

$$[B_1] = \int_{-1}^{+1}\int_{-1}^{+1} [B_q]^T[B_p]|J|\,d\eta d\zeta \tag{22b}$$

$$[K_2]_{sum} = \int_{-1}^{+1}\int_{-1}^{+1} [B_q]^T[B_q]|J|\,d\eta d\zeta. \tag{22c}$$

The matrices of Eq. (22) are determined by numerical integration. Note that $[F]$, $[B_1]$ and $[K_2]_{sum}$ depend on the discretization of the structure-medium interface and that $[F]$ and $[K_2]_{sum}$ are symmetric and semipositive definite. The latter results in $[K_0]$, $[K_1]$, and $[K_2]_{sum}$ also being symmetric. For a nontrivial nodal vector of the function $\{u\}$, the function $u = [N]^T\{u\}$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

does not vanish identically over the element. From Eqs. (15) and (22a)

$$\{u\}^T[F]\{u\} = \int_{-1}^{+1}\int_{-1}^{+1} [u \quad 0 \quad 0] \times ([J]^{-1})^T[J]^{-1}\begin{Bmatrix} u \\ 0 \\ 0 \end{Bmatrix}|J|\,d\eta d\zeta > 0 \quad (23)$$

follows as $([J]^{-1})^T[J]^{-1}|J|$ is positive definite. $[F]$ is thus positive definite.

The mass matrix equals

$$[M] = \int_V \frac{1}{c^2}[\hat{N}]^T[\hat{N}]dV \quad (24)$$

with its submatrices decomposed with respect to the interior and exterior boundaries as ($j = i, e$; $l = i, e$)

$$[M]_{jl} = \int_V \frac{1}{c^2}[\hat{N}]_j^T[\hat{N}]_l dV = \int_{-1}^{+1}\int_{-1}^{+1}\int_{-1}^{+1} \frac{1}{c^2}[\hat{N}]_j^T[\hat{N}]_l|\hat{J}|\,d\xi d\eta d\zeta \quad (25)$$

where

$$[\hat{N}]_j = \tfrac{1}{2}(1 + \xi_j\xi)[N] \quad (j = i, e) \quad (26)$$

and

$$[N] = [N_1 \quad N_2 \quad \ldots]. \quad (27)$$

Again, the integration in the ξ-direction is performed analytically. The terms of order in w higher than 1 can be neglected when the limit $w \to 0$ is performed analytically. Thus

$$[M]_{jl} \approx w[M_2]_{jl} = \frac{w}{4}\left(1 + \frac{\xi_j\xi_l}{3}\right)[B_4] \quad (28)$$

applies with

$$[B_4] = \int_{-1}^{+1}\int_{-1}^{+1} \frac{1}{c^2}[N]^T[N]|J|\,d\eta d\zeta. \quad (29)$$

Note that $[B_4]$ is symmetric and positive definite.

The coefficient matrices $[F]$, $[B_1]$, $[K_2]_{sum}$, and $[B_4]$ associated with any type of two-dimensional surface finite element (resulting from performing the analytical limit of the three-dimensional element) correspond to those associated with the two-node line element (resulting from the four-node quadrilateral element) specified in Eqs. (33) to (36) of Song and Wolf (1995). No other modification occurs. In particular, the relationship of the dynamic-stiffness matrices of similar interfaces of the unbounded medium, the assemblage of the infinitesimal finite element cell and of the unbounded medium, the derivation of the consistent infinitesimal finite element cell equation and its time discretization remain the same. The assemblage process

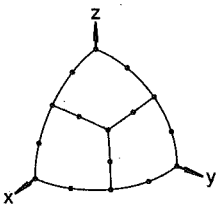


Fig. 3 Finite element mesh of structure-medium interface of spherical cavity

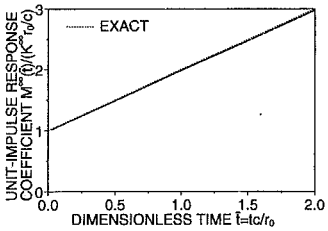


Fig. 4 Unit-impulse response coefficient for spherical cavity

of the coefficient matrices automatically incorporates the boundary conditions on interfaces extending to infinity.

Examples

Spherical Cavity With Symmetric Waves. Acoustic wave propagation resulting from a spherical cavity of radius r_0 embedded in a full space with the speed of sound c and mass density ρ is used for illustration. A uniform velocity $\bar{v}(t)$ in the normal direction of the wall, the structure-medium interface, is prescribed. For this one-dimensional problem with symmetric waves, the nodal flux-velocity potential relationship at $r = r_0$ with $R(t) = -4\pi r_0^2\bar{v}(t)$ and velocity potential $u(t)$ equals

$$R(t) = \int_0^t M^\infty(t - \tau)\ddot{u}(\tau)d\tau. \quad (30)$$

The analytical solution of the corresponding unit-impulse response coefficient equals

$$M^\infty(t) = K^\infty\left(\frac{r_0}{c} + t\right)H(t) \quad (31)$$

with the static-stiffness coefficient of the unbounded medium $K^\infty = 4\pi r_0$ and the Heaviside step function $H(t)$.

When applying the consistent infinitesimal finite element cell method, only one octant of the structure-medium interface is discretized due to symmetry (Fig. 3). Three eight-node isoparametric surface finite elements are used yielding a system with 16 degrees-of-freedom. After determining the unit-impulse response matrix of order 16×16, the uniform spatial distribution of the (second time derivative of the) velocity potential, which is selected as one, is enforced. The sum of all nodal flux corresponds to the unit-impulse response coefficient $M^\infty(t)$. The analysis is performed with $\Delta t = 0.05r_0/c$. The coefficient $M^\infty(t)$ nondimensionalized with $K^\infty r_0/c$ is plotted as a function of the dimensionless time $\bar{t} = tc/r_0$ in Fig. 4. The agreement with the exact result (Eq. (31)) is excellent.

As an application using the unit-impulse response coefficient calculated above with the consistent infinitesimal finite element cell method, the pressure response for a prescribed velocity in the normal direction is determined. Introducing the pressure $p(t) = -\rho\dot{u}(t)$ in Eq. (30) yields

$$\bar{v}(t) = \frac{1}{4\pi r_0^2\rho}\int_0^t M^\infty(t - \tau)\dot{p}(\tau)d\tau. \quad (32)$$

Discretizing Eq. (32) at time station n for a piecewise constant $M^\infty(t)$ for each time step yields

$$\bar{v}_n = \frac{1}{4\pi r_0^2\rho}\sum_{j=1}^{n} M_{n-j+1}^\infty\int_{t_{j-1}}^{t_j}\dot{p}(\tau)d\tau = \frac{1}{4\pi r_0^2\rho}\sum_{j=1}^{n} M_{n-j+1}^\infty(p_j - p_{j-1}). \quad (33)$$

p_n follows from Eq. (33) with the zero initial condition $p_0 = 0$.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

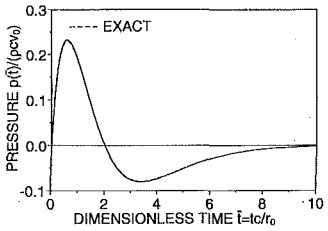


Fig. 5 Pressure response to uniform normal velocity prescribed on wall of spherical cavity

For a spherical cavity with vanishing initial pressure the following velocity in the normal direction is prescribed on its wall (amplitude v_0)

$$\bar{v}(t) = v_0 \frac{c}{r_0} t \exp\left(-\frac{c}{r_0} t\right) H(t). \tag{34}$$

The exact solution of the corresponding pressure

$$p(t) = \rho c v_0 \frac{c}{r_0} t\left(1 - \frac{1}{2}\frac{c}{r_0} t\right) \exp\left(-\frac{c}{r_0} t\right) H(t) \tag{35}$$

is easily verified by substituting Eqs. (35) and (31) in Eq. (32). The pressure determined from Eq. (33) agrees well with the exact solution (Fig. 5).

Square Prism Embedded in Inhomogeneous Half-Space. As a truly three-dimensional problem, a square prism of length $2b$ embedded with depth e in a half-space is addressed (Fig. 6(a)). The embedment ratio $e/b = \frac{2}{3}$ is selected. An inhomogeneous half-space compatible with similarity with c_1 of the unbounded medium adjacent to the walls and $c_2 = 2c_1$ adjacent to the base is calculated. The constant ρ does not vary. The structure-medium interface is first discretized by finite elements. Due to symmetry, only a quarter of the prism is analyzed. The finite element discretization of the three faces is shown in Fig. 6(b). The corresponding unit-impulse response matrix $[M^\infty(t)]$ of order 100×100 is then established with $\Delta t = 0.04b/c_1$.

To ease comparison, a uniform variation of the velocity potential (i.e., also of the pressure) over the structure-medium interface is enforced, and the resulting nodal flux are added. For a unit velocity potential, this procedure (which adds all coefficients of $[M^\infty(t)]$) leads to the equivalent unit-impulse response coefficient $M^\infty(t)$. It is plotted in Fig. 7 with $K^\infty = 17.37b$. As no analytical solution is available, an analysis with an extended mesh of finite elements is also performed. The element length in the radial direction equals $0.1b$ and $\Delta t = 0.02b/c_1$. Twenty-five cells of finite elements are necessary to prevent the reflected waves reaching the structure-medium interface for $\bar{t} = t_{c_1}/b \leq 2$. The agreement is excellent with the exception of $\bar{t} < 0.1$ where in the extended-mesh analysis the wave front which is still close to the structure-medium interface is not yet well represented (first two rows of finite elements).

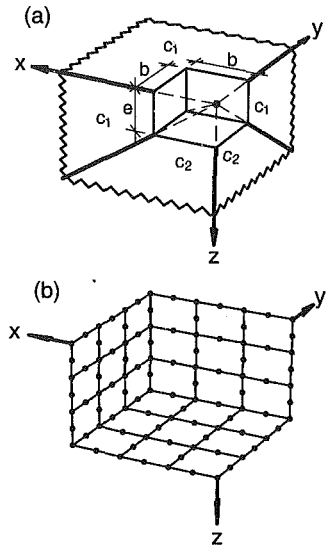


Fig. 6 One quarter of square prism embedded in inhomogeneous half-space: (*a*) geometry with center of similarity; (*b*) finite element mesh of structure-medium interface

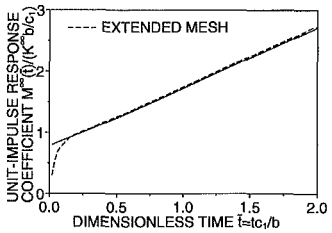


Fig. 7 Unit-impulse response coefficient for square prism embedded in inhomogeneous half-space

References

Dasgupta, G., 1982, "A Finite Element Formulation for Unbounded Homogeneous Continua," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 136–140.

Dominguez, J., 1993, *Boundary Elements in Dynamics,* Computational Mechanics Publications, Southampton, U.K.

Friedman, M. B., and Shaw, R. P., 1962, "Diffraction of a Plane Shock Wave by an Arbitrary Rigid Cylindrical Obstacle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, pp. 40–46.

Kausel, E., Roësset, J. M., and Waas, G., 1975, "Dynamic Analysis of Footings on Layered Media," ASCE *Journal of Engineering Mechanics,* Vol. 101, pp. 679–693.

Song, C., and Wolf, J. P., 1995, "Consistent Infinitesimal Finite-Element Cell Method: Out-of-Plane Motion," ASCE *Journal of Engineering Mechanics,* Vol. 121, pp. 613–619.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Wrinkling of Rectangular Webs Under Nonlinearly Distributed Edge Loading[1]

C. C. Lin
Design Engineer,
Quantum Corporation,
500 McCarthy Blvd.,
Milpitas, CA 95035
Assoc. Mem. ASME

C. D. Mote, Jr.
Vice Chancellor and FANUC Chair
in Mechanical Systems,
Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720
Fellow ASME

Nonlinearly distributed in-plane tension at the edges of a thin, flat, rectangular material with small flexural stiffness, called a web, can cause it to wrinkle. Airy stress functions are determined in closed form for webs under arbitrarily prescribed, in-plane edge loading. The loading at impending wrinkling, and the corresponding wrinkling shape in the web, are predicted by the solution of an eigenvalue problem. This analysis is illustrated through two examples: (I) wrinkling of a web induced by nearly uniform normal edge tension; (II) wrinkling of a web induced by uniform normal tension distributed over a section of the edge. Predictions of wrinkling using membrane theory are compared to those obtained with the current web theory. The contribution of small flexural stiffness to the prediction of the condition of wrinkling is shown to be substantial.

Introduction

Thin, flat, rectangular materials with small flexural stiffness, generally referred to as webs, appear as paper webs, plastic sheets, films, magnetic media, textiles, and the like. Wrinkling of webs during handling processes is a common problem limiting the technology. Wrinkling degrades the web quality, increases defects, causes machine downtime, and can lead to rejection of web products. To improve web handling machine performance and to reduce the losses resulting from wrinkling, the causes of wrinkling must be identified and controlled.

The deflection of the web under aerodynamic transverse loading is usually large compared to its thickness, and the in-plane compression and shear that are induced by deformation can be significant to wrinkling (Lin and Mote, 1995). A web may be placed in simple shear during handling and the web wrinkles when the shear exceeds a critical value. In this case, the region of the web surface that wrinkles is contained within a predictable parallelogram (Lin and Mote, 1996).

Reissner (1938) showed that a wrinkled membrane (zero flexural stiffness) is stressed in tension along the direction of the wrinkle line and in compression perpendicular to the wrinkle line. When the state of deformation is prescribed, the directions of the wrinkle lines can be determined from the elements of the Green's deformation tensor (Pipkin, 1986; Lin and Mote, 1996). Roddeman et al. (1987) defined three stress states for membranes: taut, slack, or wrinkled. Compression in a membrane always causes wrinkling (Steigmann and Pipkin, 1989). However, a web with small flexural stiffness can support nonvanishing in-plane compressive stress of magnitude less than a critical value determined by the flexural stiffness. If one in-plane principal stress is tensile and the other is compressive with magnitude greater than the critical value, the web wrinkles (Lin and Mote, 1996).

During manufacturing processes, an unsupported web travels under a nearly uniform, longitudinal edge tension between the rollers. The tension can be represented as the superposition of a uniform tension and a small (perturbed), nonuniform normal edge loading. Compression can be generated in the interior region of the web when the edge tension is nonlinearly distributed. Wrinkle lines oriented parallel to the longitudinal coordinate may appear if the magnitude of the perturbed, nonlinearly distributed, longitudinal edge loading exceeds a critical value. The size and the shape of the wrinkled region, and the directions of the wrinkle lines of the web under nonlinearly distributed edge loading, can be quite different from those found in a web under uniform compression and under simple shear.

The purpose of this paper is to predict the wrinkling of thin, flat, rectangular webs under nonlinearly distributed, in-plane edge loading. Gorman and Singhal's method (1993) is applied to determine the Airy stress functions for webs under in-plane edge loading whose distributions are described by sine functions. Superposition of those Airy functions permits the representation of the stress field generated by arbitrarily prescribed edge loading. The loading at impending wrinkling and the corresponding wrinkling shape in the web are predicted by the solution of an eigenvalue problem.

The analysis is illustrated through two examples: (I) wrinkling of a web induced by nearly uniform normal edge tension; (II) wrinkling of a web induced by uniform normal tension distributed over a section of the edge. Predictions of wrinkling using membrane theory are contrasted with those obtained using the current web theory. Small flexural stiffness is shown to significantly effect the prediction of wrinkling and the wrinkling shape.

Stress Field Induced by Edge Loading

Consider a thin, flat, rectangular web under normal in-plane loading along opposite edges (see Fig. 1). Let $x = \hat{x}/L$ and $y = \hat{y}/L$ be the dimensionless in-plane coordinates and $s = B/L$ be the aspect ratio of the web. The Airy stress function F describing a linearly distributed longitudinal tension (Fig. 1(a, b)) is

$$F = A_1 y^2 + A_2 y^3, \tag{1}$$

where A_1 and A_2 are constants. Width compression and shear will not appear in the web because $F_{,xx} = 0$ and $F_{,xy} = 0$. (Comma denotes partial differentiation.)

Gorman and Singhal's method (1993) is applied to determine the Airy stress function for the web under loading T_{xx} along

[1] The authors would like to thank the James River Corporation and the TAPPI Research Foundation for their support and interest in this work.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 2, 1995; final revision, Oct. 18, 1995. Associate Technical Editor: R. Abeyaratne.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

edges $x = 0$ and $x = 1$. The distribution of the loading is described by a sine series:

$$T_{xx}(0, y) = T_{xx}(1, y) = \sum_{m=1}^{\tilde{m}} \lambda_m \sin\left(\frac{m\pi}{s} y\right), \quad (2)$$

where $\tilde{m} \to \infty$. For the edge loading symmetric with respect to axis $y = s/2$ (Fig. 1(*c*)), m is an infinite sequence of odd integers greater than zero; and for the antisymmetric edge loading, m is a sequence of even integers. The Airy function under the symmetric edge loading is

$$\begin{aligned} F = &\left\{ \sum_{m=1,3}^{\tilde{m}} \lambda_m \sin\left(\frac{m\pi}{s} y\right)\left[C_{1m} \cosh\left(\frac{m\pi}{s}\left(x - \frac{1}{2}\right)\right)\right.\right. \\ &\left.\left. + C_{2m}\left(x - \frac{1}{2}\right) \sinh\left(\frac{m\pi}{s}\left(x - \frac{1}{2}\right)\right)\right]\right\} \\ &+ \left\{ \sum_{p=1,3}^{\tilde{p}} B_p \sin(p\pi x)\left[D_{1p} \cosh\left(p\pi\left(y - \frac{s}{2}\right)\right)\right.\right. \\ &\left.\left. + D_{2p}\left(y - \frac{s}{2}\right) \sinh\left(p\pi\left(y - \frac{s}{2}\right)\right)\right]\right\} \\ &+ \left\{ \sum_{q=1,3}^{\tilde{q}} \bar{B}_q \sin\left(\frac{q\pi}{s} y\right)\left[\bar{D}_{1q} \cosh\left(\frac{q\pi}{s}\left(x - \frac{1}{2}\right)\right)\right.\right. \\ &\left.\left. + \bar{D}_{2q}\left(x - \frac{1}{2}\right) \sinh\left(\frac{q\pi}{s}\left(x - \frac{1}{2}\right)\right)\right]\right\}, \quad (3) \end{aligned}$$

where

$$\tilde{p}, \tilde{q} \to \infty \quad (4a)$$

$$C_{1m} = -\frac{2s^2\left[\sinh\left(\frac{m\pi}{2s}\right) + \frac{m\pi}{2s}\cosh\left(\frac{m\pi}{2s}\right)\right]}{m^2\pi^2\left[\frac{m\pi}{s} + \sinh\left(\frac{m\pi}{s}\right)\right]} \quad (4b)$$

$$C_{2m} = \frac{2s \sinh\left(\frac{m\pi}{2s}\right)}{m\pi\left[\frac{m\pi}{s} + \sinh\left(\frac{m\pi}{s}\right)\right]} \quad (4c)$$

$$D_{1p} = \frac{s \sinh\left(\frac{p\pi s}{2}\right)}{p\pi[p\pi s + \sinh(p\pi s)]} \quad (4d)$$

$$D_{2p} = -\frac{2 \cosh\left(\frac{p\pi s}{2}\right)}{p\pi[p\pi s + \sinh(p\pi s)]} \quad (4e)$$

$$\bar{D}_{1q} = \frac{s \sinh\left(\frac{q\pi}{2s}\right)}{q\pi\left[\frac{q\pi}{s} + \sinh\left(\frac{q\pi}{s}\right)\right]} \quad (4f)$$

$$\bar{D}_{2q} = -\frac{2s \cosh\left(\frac{q\pi}{2s}\right)}{q\pi\left[\frac{q\pi}{s} + \sinh\left(\frac{q\pi}{s}\right)\right]}. \quad (4g)$$

The expression within each brace on the right-hand side of (3) satisfies $\nabla^4\{\ \} = 0$. The expression in the first brace on the right-hand side of (3), satisfies the normal stress boundary condition. The shear stress boundary condition is satisfied by adding the terms in the second and the third braces to those in the first brace. B_p and $\bar{B}_q$ are chosen to ensure vanishing shear stress along the boundaries.

The Airy function under the antisymmetric edge loading is

$$\begin{aligned} F = &\left\{ \sum_{m=2,4}^{\tilde{m}} \lambda_m \sin\left(\frac{m\pi}{s} y\right)\left[C_{1m} \cosh\left(\frac{m\pi}{s}\left(x - \frac{1}{2}\right)\right)\right.\right. \\ &\left.\left. + C_{2m}\left(x - \frac{1}{2}\right) \sinh\left(\frac{m\pi}{s}\left(x - \frac{1}{2}\right)\right)\right]\right\} \\ &+ \left\{ \sum_{p=1,3}^{\tilde{p}} \hat{B}_p \sin(p\pi x)\left[\hat{D}_{1p} \sinh\left(p\pi\left(y - \frac{s}{2}\right)\right)\right.\right. \\ &\left.\left. + \hat{D}_{2p}\left(y - \frac{s}{2}\right) \cosh\left(p\pi\left(y - \frac{s}{2}\right)\right)\right]\right\} \\ &+ \left\{ \sum_{q=2,4}^{\tilde{q}} \bar{B}_q \sin\left(\frac{q\pi}{s} y\right)\left[\bar{D}_{1q} \cosh\left(\frac{q\pi}{s}\left(x - \frac{1}{2}\right)\right)\right.\right. \\ &\left.\left. + \bar{D}_{2q}\left(x - \frac{1}{2}\right) \sinh\left(\frac{q\pi}{s}\left(x - \frac{1}{2}\right)\right)\right]\right\}, \quad (5) \end{aligned}$$

where $\hat{B}_p$ and $\bar{B}_q$ are chosen to ensure vanishing shear stress along the boundaries, and

$$\hat{D}_{1p} = \frac{s \cosh\left(\frac{p\pi s}{2}\right)}{p\pi[-p\pi s + \sinh(p\pi s)]} \quad (6a)$$

$$\hat{D}_{2p} = -\frac{2 \sinh\left(\frac{p\pi s}{2}\right)}{p\pi[-p\pi s + \sinh(p\pi s)]}. \quad (6b)$$

The stress state in the web from (3) under symmetric edge loading along $x = 0$ and $x = 1$ with $s = 0.5$, $\lambda_1 = 1$, and all other $\lambda_m = 0$ is shown in Fig. 2. The normal stresses T_{xx} in Fig. 2(*a*) are nearly uniform away from the edges. The shear stresses T_{xy} in Fig. 2(*b*) are antisymmetric about $x = \frac{1}{2}$ and $y = s/2$. The normal stresses T_{yy} in Fig. 2(*c*, *d*) become compressive in the interior of the web. The relative variation between the compressive stress of largest magnitude in the width direction given by (3) when $\tilde{p} = \tilde{q} = 5$ and that given when $\tilde{p} = \tilde{q} = 41$ is less than 0.6 percent.

Because all prescribed longitudinal in-plane edge loading can be modeled by symmetric and antisymmetric components, the stress distribution in the web can be determined from a superposition of the corresponding Airy functions in (3) and (5).

Wrinkling Induced by Nearly Uniform Edge Tension

During manufacturing processes, an unsupported web may travel under a nearly uniform, longitudinal edge tension between the rollers. The tension can be modeled as the superposition of a uniform tension of magnitude 1 and a relatively small, nonlinearly distributed component T_{xx} in (2).

The dimensionless equation of equilibrium of the web is

$$-w_{,xx} - (F_{,yy}w_{,xx} - 2F_{,xy}w_{,xy} + F_{,xx}w_{,yy}) + \epsilon\nabla^4 w = 0, \quad (7)$$

where

$$w = \frac{\hat{w}}{h}, \quad F = \frac{\hat{F}}{TL^2}, \quad \epsilon = \frac{D}{TL^2}. \quad (8)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

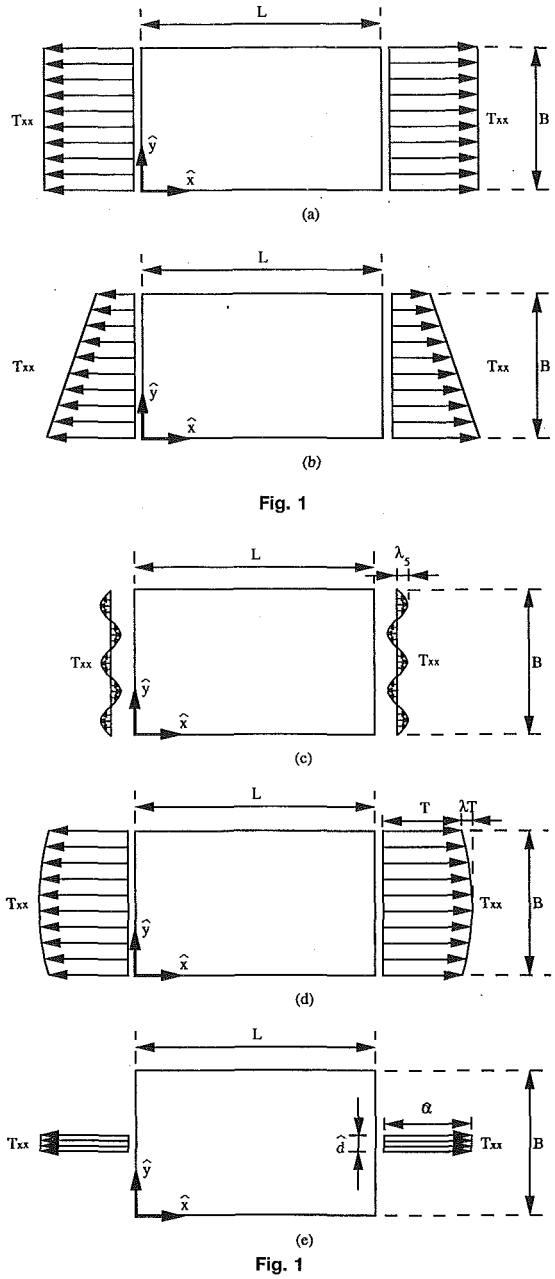


Fig. 1 Rectangular webs under longitudinal loading along two opposite edges; (*a*) uniform tension, (*b*) linearly distributed tension, (*c*) sine distribution (all $\lambda_m = 0$ except λ_5), (*d*) nearly uniform tension, and (*e*) uniform tension over a section of the edge

$\hat{w}(\hat{x}, \hat{y})$ denotes the transverse displacement; F is the Airy stress function, derived from the small nonlinearly distributed component T_{xx} in (2); $D = Eh^3/[12(1 - \nu^2)]$; E is Young's modulus; h is thickness; ν is Poisson's ratio. The material properties are uniform and constant. The stiffness ratio, ϵ, equals the ratio of the flexural stiffness to the stiffness derived from the uniform longitudinal tension and is normally small.

When the web is under nonlinearly distributed edge loading, width compression can be generated in the interior region of the web. The magnitude of the nonlinearly distributed component T_{xx} in (2) at impending wrinkling and the corresponding wrinkling mode of the web are predicted from the solution of an eigenvalue problem derived from (7) plus appropriate boundary conditions. For example, consider the web in Fig. 1(*d*) under edge tension $T[1 + \lambda \sin(\pi y/s)]$. In this case, $\lambda_1 = \lambda$ and all other $\lambda_m = 0$ in (3). The eigenvalue problem, formed

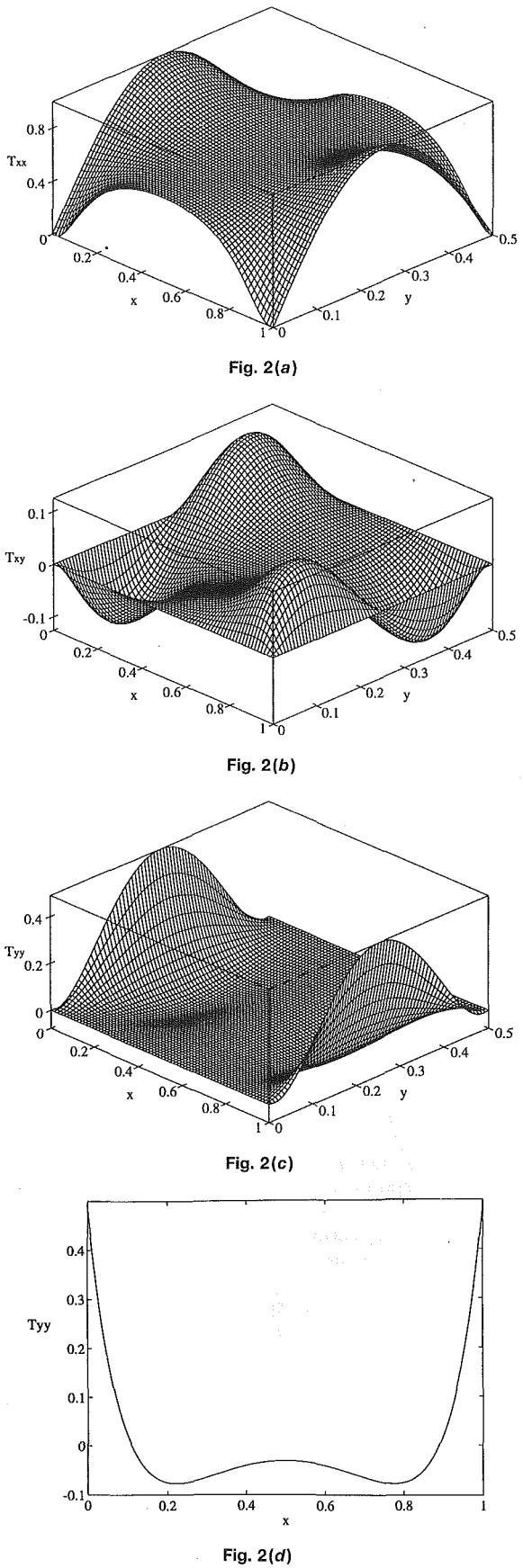


Fig. 2 The stress state in the web from (3) under symmetric edge loading (2) along $x = 0$ and $x = 1$ with $s = 0.5$, $\lambda_1 = 1$, and all other $\lambda_m = 0$; (*a*) the normal stresses in the longitudinal direction ($T_{xx} = F_{,yy}$), (*b*) the shear stresses ($T_{xy} = -F_{,xy}$), (*c*) the normal stresses in the width direction ($T_{yy} = F_{,xx}$), and (*d*) the stresses T_{yy} along $y = s/2$ in Fig. 2(*c*)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

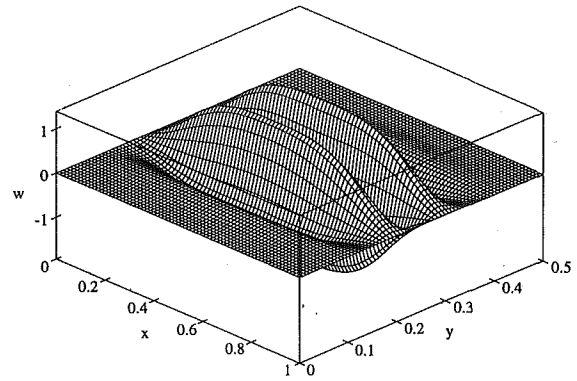


Fig. 3(*a*)

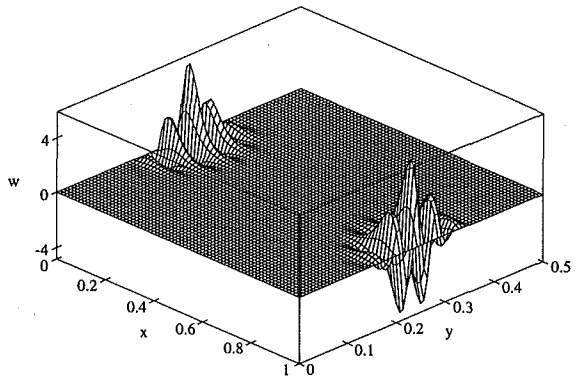


Fig. 3(*b*)

Fig. 3 The wrinkling shapes of the web under the edge loading in Fig. 1(*d*) with $\epsilon = 10^{-6}$, $s = 0.5$, and (*a*) $\lambda_{cr}^{P} = 0.1566$ and (*b*) $\lambda_{cr}^{N} = -0.2643$

by (7) plus simple support boundary conditions, is solved using the Galerkin method with comparison functions w_{ij},

$$w_{ij} = \sin(i\pi x) \sin\left(\frac{j\pi y}{s}\right), \tag{9}$$

where $i, j = 1, 2, 3, \ldots$. When $\epsilon = 10^{-6}$ and $s = 0.5$, the smallest positive eigenvalue and the largest negative eigenvalue are $\lambda_{cr}^{P} = 0.1566$ and $\lambda_{cr}^{N} = -0.2643$. The relative variation between λ_{cr}^{P} predicted by F in (3) when $\tilde{p} = \tilde{q} = 5$ and that predicted when $\tilde{p} = \tilde{q} = 41$ is less than 0.03 percent, and this variation is less than 0.5 percent for λ_{cr}^{N}. The web with small stiffness can support compressive stress of magnitude less than some critical value without wrinkling. The magnitude of the width compression induced in the web increases as $|\lambda|$ increases. This web wrinkles when $\lambda \geq \lambda_{cr}^{P}$ or $\lambda \leq \lambda_{cr}^{N}$ and is smooth when $\lambda_{cr}^{N} < \lambda < \lambda_{cr}^{P}$.

The wrinkling mode corresponding to $\lambda = \lambda_{cr}^{P} = 0.1566$ is shown in Fig. 3(*a*). The dimensionless stresses in the web are λ_{cr}^{P} times the stresses in Fig. 2 plus uniform longitudinal tension of magnitude 1,

$$T_{xx} = 1 + \lambda_{cr}^{P} T_{xx}^{(2)} \tag{10a}$$

$$T_{xy} = \lambda_{cr}^{P} T_{xy}^{(2)} \tag{10b}$$

$$T_{yy} = \lambda_{cr}^{P} T_{yy}^{(2)}, \tag{10c}$$

where $T_{xx}^{(2)}$, $T_{xy}^{(2)}$, and $T_{yy}^{(2)}$ are the stresses shown in Fig. 2. The directions of the wrinkle lines are parallel to the longitudinal tension and perpendicular to the width compression appearing in the interior region of the web. Shear stresses are anti-symmetric with respect to the axes $x = \frac{1}{2}$ and $y = s/2$ and do not effect the directions of the wrinkle lines. The wrinkles extend the length of the web though $T_{yy} > 0$ in the regions near the edges $x = 0$ and $x = 1$.

The wrinkling mode corresponding to $\lambda = \lambda_{cr}^{N} = -0.2643$ is shown in Fig. 3(*b*). For negative λ in Fig. 1(*d*), the applied longitudinal tension near the middle of the edges $x = 0$ and $x = 1$ is smaller than that near the ends at $y = 0$ and $y = 0.5$ of those edges. The dimensionless stresses in the web are

$$T_{xx} = 1 + \lambda_{cr}^{N} T_{xx}^{(2)} \tag{11a}$$

$$T_{xy} = \lambda_{cr}^{N} T_{xy}^{(2)} \tag{11b}$$

$$T_{yy} = \lambda_{cr}^{N} T_{yy}^{(2)}. \tag{11c}$$

The web is under width compression in the regions near the edges $x = 0$ and $x = 1$ and wrinkles near the middle of those edges. The shape and the size of the wrinkled regions are different from those of the regions under width compression. The directions of the wrinkle lines are parallel to the longitudinal axis. The antisymmetric shear stresses do not effect the directions of the wrinkle lines.

Wrinkling Induced by Nonlinearly Distributed Edge Loading

The dimensionless equation of equilibrium of the web under nonlinearly distributed, longitudinal in-plane edge loading described in (2) is

$$\nabla^4 w - (F_{,yy} w_{,xx} - 2F_{,xy} w_{,xy} + F_{,xx} w_{,yy}) = 0. \tag{12}$$

Without uniform longitudinal tension T applied at the boundaries as in (7), F in (12) is nondimensionalized by D,

$$F = \frac{\hat{F}}{D}. \tag{13}$$

The magnitude of the edge loading at impending wrinkling and the corresponding wrinkling modes of the web are predicted from the solution of an eigenvalue problem derived from (12) plus appropriate boundary conditions. For example, for the web in Fig. 1(*e*), the loading $\hat{\alpha}$ is uniformly distributed over width $\hat{d}$. $F_{,xx}$, $F_{,xy}$, and $F_{,yy}$ vanish along the edges of the web except for

$$F_{,yy}(0, y) = F_{,yy}(1, y) = \alpha[u(y - 0.5(s - d)) - u(y - 0.5(s + d))], \tag{14}$$

where $u(y - a)$ is a unit step distribution initiated at $y = a$. The dimensionless parameters are

$$\alpha = \hat{\alpha} \frac{L^2}{D}, \quad d = \frac{\hat{d}}{s}. \tag{15}$$

The right-hand side of (14) can be represented in a Fourier sine series (2).

The eigenvalue problem, derived from (12) and (14) plus simple support boundary conditions, is solved using the Galerkin method. For the web with $s = 0.5$ and $d = 0.025$, the smallest positive eigenvalue and the largest negative eigenvalue are $\alpha_{cr}^{P} = 127613$ and $\alpha_{cr}^{N} = -4941.34$. The web wrinkles when $\alpha \geq \alpha_{cr}^{P}$ or $\alpha \leq \alpha_{cr}^{N}$ and is smooth when $\alpha_{cr}^{N} < \alpha < \alpha_{cr}^{P}$. The relative variation between α_{cr}^{N} predicted by F in (3) when $\tilde{p} = \tilde{q} = 5$ and $\tilde{m} = 15$ and that predicted when $\tilde{p} = \tilde{q} = 21$ and $\tilde{m} = 51$ is less than 0.8 percent, and the variation is less than 7 percent for α_{cr}^{P}.

The magnitude of the tension α_{cr}^{P} is typically 10 to 100 times that of the compression $|\alpha_{cr}^{N}|$ required to wrinkle the web. For the web with $s = 0.5$ and $d = 0.025$, $|\alpha_{cr}^{P}/\alpha_{cr}^{N}| = 26$.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

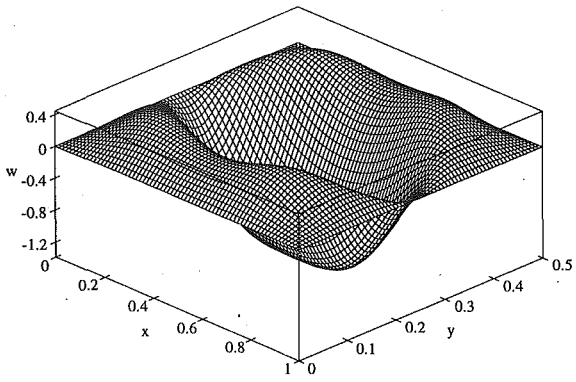


Fig. 4(*a*)

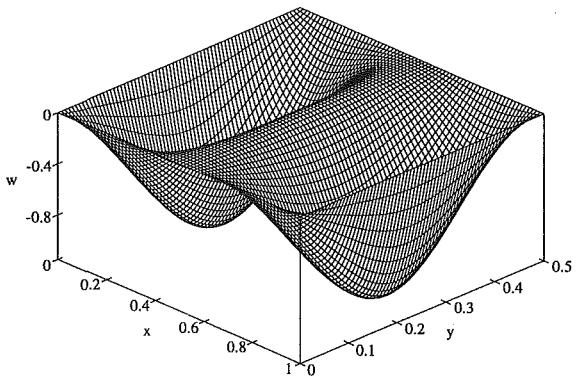


Fig. 4(*b*)

Fig. 4 The wrinkling shapes of the web under the edge loading in Fig. 1(*e*) with $s = 0.5$, $d = 0.025$, and (*a*) $\alpha_{cr}^{P} = 127613$ and (*b*) $\alpha_{cr}^{N} = -4941.34$

The wrinkling mode corresponding to $\alpha_{cr}^{P} = 127613$ is shown in Fig. 4(*a*). The direction of the wrinkle line is perpendicular to the width compression appearing in the middle of the web. A similar wrinkling mode is observed when the middle of opposite edges of a piece of rectangular paper web are pinched by fingers and pulled.

The wrinkling mode corresponding to $\alpha_{cr}^{N} = -4941.34$ is shown in Fig. 4(*b*). For negative α_{cr}^{N} in Fig. 1(*e*), the web is under uniformly distributed compression over width $\hat{d}$ in the middle of the longitudinal edges. The directions of the wrinkle lines are perpendicular to the longitudinal compression.

Discussion

A web modeled by membrane theory wrinkles under arbitrarily small nonvanishing compressive stress. Membrane theory underestimates the magnitude of the loading at impending wrinkling. For instance, membrane theory predicts that any λ in Fig. 1(*d*) and/or α in (14), with arbitrarily small magnitude, wrinkles the web. However, a web with small flexural stiffness is smooth when $\lambda_{cr}^{N} < \lambda < \lambda_{cr}^{P}$ or $\alpha_{cr}^{N} < \alpha < \alpha_{cr}^{P}$. Membrane theory predicts that the wrinkled regions are those under compression and also that an infinite number of wrinkle lines appear in the wrinkled regions. However, the wrinkles extend the length of the web even though $T_{xx} > 0$ and $T_{yy} > 0$ in the regions near the edges $x = 0$ and $x = 1$ ((10) and Fig. 3(*a*)). The shapes and the sizes of the wrinkled regions are different from those of the regions under width compression ((11) and Fig. 3(*b*)). A finite number of wrinkle lines appears in the wrinkled region in Fig. 3. Thus, the effects of small flexural stiffness on the prediction of the critical nonlinear component of the edge loading that causes wrinkling, on the number and the direction of the wrinkle lines, and on the sizes and shapes of the wrinkled regions are significant and can not be neglected.

Conclusions

1 Linearly distributed normal tension, applied along two opposite edges or applied along all four edges of a rectangular web will not cause wrinkling.
2 When the magnitude of the nonlinear component of the edge loading is greater than a critical value, the web wrinkles and the number of wrinkle lines in the web is finite. The critical value and the corresponding wrinkling shape of the web can be predicted from the solution of an eigenvalue problem.
3 At impending wrinkling, the critical magnitudes of the nonlinear component of the nearly uniform edge tension given by the smallest positive and the largest negative eigenvalues are of the same order. The critical magnitude of the tension under uniform normal loading distributed over a section of the edge is typically ten to one hundred times larger than the critical compression required to wrinkle the web.
4 The effects of small flexural stiffness on the prediction of the critical nonlinear component of the edge loading that causes wrinkling, on the number and the direction of the wrinkle lines, and on the sizes and shapes of the wrinkled regions are significant and can not be neglected.

References

Gorman, D. J., and Singhal, R. K., 1993, "A Superposition-Rayleigh-Ritz Method for Free Vibration Analysis of Non-Uniformly Tensioned Membranes," *Journal of Sound and Vibration,* Vol. 162, No. 3, pp. 489–501.

Lin, C. C., and Mote, C. D., Jr., 1995, "Equilibrium Displacement and Stress Distribution in a Two Dimensional, Axially Moving Web Under Transverse Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 62, pp. 772–779.

Lin, C. C., and Mote, C. D., Jr., 1996, "The Wrinkling of Thin, Flat, Rectangular Webs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 774–779.

Pipkin, A. C., 1986, "The Relaxed Energy Density for Isotropic Elastic Membranes," *IMA Journal of Applied Mathematics,* Vol. 36, pp. 85–99.

Reissner, E., 1938, "On Tension Field Theory," *Proceedings of the Fifth International Congress for Applied Mechanics,* pp. 88–92.

Roddeman, D. G., Drukker, J., Oomens, C. W. J., and Janssen, J. D., 1987, "The Wrinkling of Thin Membranes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 884–887.

Steigmann, D. J., and Pipkin, A. C., 1989, "Wrinkling of Pressurized Membranes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 624–628.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

H. Murakami
Professor,
Mem. ASME.

E. Reissner
Professor Emeritus,
Honorary Mem. ASME.

J. Yamakawa
Graduate Student,
Assoc. Mem. ASME.

Department of Applied Mechanics
and Engineering Sciences,
University of California at San Diego,
La Jolla, CA 92093-0411

Anisotropic Beam Theories With Shear Deformation

We investigate the effect of constitutive coupling of stretching, bending, and transverse shearing deformation on the deflection of an anisotropic cantilever beam with narrow rectangular cross-section. To this end, we have developed a hierarchy of beam models by applying a variational principle for displacements and transverse *stresses to the associated plane stress problem.*

Introduction

As a step towards the analysis of beam, plate, and shell problems with constitutive coupling of stretching, bending, and transverse shearing, we consider in what follows an anisotropic plane stress problem for a rectangular region with conditions of loading and support corresponding to those of a cantilever beam.

In order to accomplish the above objective in such a way as to be applicable to laminated as well as to homogeneous construction our method of solution makes use of a variational principle for displacements and *transverse* stresses (Reissner, 1984). This approach is intended to allow in a variational way the satisfaction of continuity conditions for displacements and transverse stresses without also requiring continuity for spanwise stresses in as much as these may have to be discontinuous to account for lamination effects. The above procedure has previously been used for the derivation of two-dimensional theories for given three-dimensional problems for *orthotropic* rather than anisotropic materials (Reissner, 1984). However, this earlier work avoids the consideration of the kind of constitutive coupling which is of particular interest in this paper. An alternate approach to the solution of our problem, involving double power series expressions for the solution of the relevant Airy equation, has earlier been presented by Hashin (1967).

In order to understand the deformation of orthotropic cantilever beams analytical solutions were obtained both for the limiting cases of infinite transverse Young's modulus $E_{yy} \to \infty$ (Hildebrand, 1943), and zero transverse Young's modulus $E_{yy} = 0$ (Reissner, 1940). Upper bounds and lower bounds of the compliance, i.e., the transverse end displacement due to a unit transverse end load, were also obtained to clarify the effects of orthotropy and transverse shear deformation (Nair and Reissner, 1975).

The problem of anisotropic coupling, in which normal strains are caused not only by normal stresses but also by shearing stresses, and shearing strains are caused not only by shearing stresses but also by normal stresses, was investigated by Nair and Reissner (1976) only for the limiting case of $E_{yy} \to \infty$. In what follows, we investigate the effect of constitutive coupling for general anisotropic beams.

Formulation of the Problem

The differential equations of the plane stress problem consist of equilibrium equations

$$\sigma_{x,x} + \tau_{,y} = 0, \quad \tau_{,x} + \sigma_{y,y} = 0, \qquad (1a, b)$$

in conjunction with strain-displacement relations

$$\epsilon_x = u_{,x}, \quad \epsilon_y = v_{,y}, \quad \gamma = u_{,y} + v_{,x}, \qquad (2a, b, c)$$

and constitutive relations

$$\epsilon_x = \frac{\sigma_x}{E_{xx}} + \frac{\sigma_y}{E_{xy}} + \frac{\tau}{G_x},$$

$$\epsilon_y = \frac{\sigma_x}{E_{yx}} + \frac{\sigma_y}{E_{yy}} + \frac{\tau}{G_y},$$

$$\gamma = \frac{\sigma_x}{G_x} + \frac{\sigma_y}{G_y} + \frac{\tau}{G}. \qquad (3)$$

The coefficients E and G in (3) are constants or given functions of x and y, with $E_{xy} = E_{yx}$. In what follows the system (3) will be used in the semi-inverted form

$$\sigma_x = K_{\sigma\epsilon}\epsilon_x + K_{\sigma\sigma}\sigma_y + K_{\sigma\tau}\tau,$$

$$\epsilon_y = -K_{\sigma\sigma}\epsilon_x + K_{\epsilon\sigma}\sigma_y + K_{\epsilon\tau}\tau,$$

$$\gamma = -K_{\sigma\tau}\epsilon_x + K_{\epsilon\tau}\sigma_y + K_{\gamma\tau}\tau, \qquad (4)$$

where

$$K_{\sigma\epsilon} = E_{xx}, \quad K_{\sigma\sigma} = -\frac{E_{xx}}{E_{xy}}, \quad K_{\sigma\tau} = -\frac{E_{xx}}{G_x},$$

$$K_{\epsilon\sigma} = \frac{1}{E_{yy}} - \frac{E_{xx}}{E_{xy}^2}, \quad K_{\epsilon\tau} = \frac{1}{G_y} - \frac{E_{xx}}{G_x E_{xy}},$$

$$K_{\gamma\tau} = \frac{1}{G} - \frac{E_{xx}}{G_x^2}. \qquad (5)$$

The notation in (4) is chosen so as to be compatible with the variational statement of the boundary value problem, which is to follow.

The system (1) to (3) is here associated with the following boundary conditions:

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 9, 1995; final revision, Aug. 8, 1995. Associate Technical Editor: S. Kryiakides.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\sigma_y = \tau = 0 \quad \text{at} \quad y = \pm c, \tag{6a}$$

$$u = 0, \quad v = 0 \quad \text{at} \quad x = 0, \tag{6b}$$

$$u = \hat{u}(y), \quad v = \hat{v}(y) \quad \text{at} \quad x = L, \tag{6c}$$

with $\hat{u}$ and $\hat{v}$ being prescribed functions of y.

In order to subsume the above boundary value problem we employ the variational principle for displacements and *transverse* stresses (Reissner, 1984). It is observed that the semi-inverted constitutive relations (4) are associated with a "semi-complementary" energy function W of the form

$$W = \tfrac{1}{2} K_{\sigma\epsilon}\epsilon_x^2 + K_{\sigma\sigma}\epsilon_x\sigma_y + K_{\sigma\tau}\epsilon_x\tau - \tfrac{1}{2} K_{\epsilon\sigma}\sigma_y^2 - K_{\epsilon\tau}\sigma_y\tau - \tfrac{1}{2} K_{\gamma\tau}\tau^2, \tag{7}$$

in terms of which

$$\sigma_x = \frac{\partial W}{\partial \epsilon_x}, \quad \epsilon_y = -\frac{\partial W}{\partial \sigma_y}, \quad \gamma = -\frac{\partial W}{\partial \tau}. \tag{8a,b,c}$$

The variational statement of the given boundary value problem is with (2), (6a,b,c), and (8a) as constraint conditions

$$\delta \iint [W(u_{,x}, \sigma_y, \tau) + v_{,y}\sigma_y + (u_y + v_{,x})\tau]dxdy = 0. \tag{9}$$

Upon carrying out the indicated variations and upon observing the indicated constraint relations, we obtain from (9), with the help of integrations by parts the developed form,

$$\iint \left[\left(v_{,y} + \frac{\partial W}{\partial \sigma_y} \right)\delta\sigma_y + \left(u_{,y} + v_{,x} + \frac{\partial W}{\partial \tau} \right)\delta\tau - \left\{ \left(\frac{\partial W}{\partial u_{,x}} \right)_{,x} + \tau_{,y} \right\}\delta u - (\tau_{,x} + \sigma_{y,y})\delta v \right] dxdy = 0. \tag{10}$$

The Euler equations of (10) are, as they should be, (1a,b), and (8b,c).

Derivation of Approximate One-Dimensional Formulations

In order to derive a system of one dimensional differential equations from (10), we approximate the axial displacement u and the transverse stresses, τ and σ_y, in the form:

$$u(x, \eta) = \sum_{j=0}^{J} U_j(x) f_j(\eta), \tag{11}$$

$$c\tau(x, \eta) = \sum_{j=1}^{J} Q_j(x) g_j^{\cdot}(\eta), \tag{12a}$$

$$\sigma_y(x, \eta) = \sum_{j=2}^{J} S_j(x) g_j(\eta), \tag{12b}$$

where $\eta = y/c$, and

$$g_j^{\cdot}(\pm 1) = g_j(\pm 1) = 0. \tag{12c}$$

In what follows, we will omit summation symbols with the repetition of subscripts indicating summation.

For present purpose we assume that the functions f_j are Legendre polynomials, i.e.,

$$f_0 = 1, \quad f_1 = \eta, \quad f_2 = \tfrac{3}{2}\eta^2 - \tfrac{1}{2}, \quad f_3 = \tfrac{5}{2}\eta^3 - \tfrac{3}{2}\eta, \ldots \tag{13}$$

and the functions $g_j^{\cdot}$ and g_j are assumed to be of the form

$$g_1^{\cdot} = \tfrac{3}{4}(1 - \eta^2), \quad g_2^{\cdot} = \eta - \eta^3,$$

$$g_3^{\cdot} = 1 - 6\eta^2 + 5\eta^4, \ldots, \tag{14a}$$

$$g_2 = \frac{-1}{4} + \frac{1}{2}\eta^2 - \frac{1}{4}\eta^4, \quad g_3 = \eta - 2\eta^3 + \eta^5, \ldots, \tag{14b}$$

where we note that

$$g_2^{\cdot} = \frac{dg_2}{d\eta}, \quad g_3^{\cdot} = \frac{dg_3}{d\eta}, \quad \ldots. \tag{14c}$$

With the relations between g_j and $g_j^{\cdot}$ as in (14) it is not necessary to make assumptions concerning the form of the transverse displacement v. Instead the result to be obtained involves weighted averages of v, of the form

$$V_j(x) = \int_{-1}^{+1} v(x, \eta) g_j^{\cdot}(\eta) d\eta. \tag{15}$$

The introduction of $\delta\sigma_y$, $\delta\tau$, δu, and δv from (11), (12a,b), and (15) into (10) gives

$$\int_0^L \left[\left(c\int_{-1}^{+1} \frac{\partial W}{\partial \sigma_y} g_j d\eta - V_j \right)\delta S_j + \left\{ \int_{-1}^{+1} \left(u_{,y} + \frac{\partial W}{\partial \tau} \right) g_j^{\cdot} d\eta + \mathrm{V}_{j,x} \right\}\delta Q_j - c\int_{-1}^{+1} \left\{ \left(\frac{\partial W}{\partial u_{,x}} \right)_{,x} + \tau_{,y} \right\} f_j d\eta \, \delta U_j - (Q_{j,x} + S_j)\delta V_j \right] dx = 0. \tag{16}$$

With σ_y, τ, and u as in (12a,b) and (11) and W as in (7) we obtain from (16) the one-dimensional system of differential equations.

$$cU_{m,x}\int_{-1}^{+1} K_{\sigma\sigma} f_m g_j d\eta - cS_m \int_{-1}^{+1} K_{\epsilon\sigma} g_m g_j d\eta - Q_m \int_{-1}^{+1} K_{\epsilon\tau} g_m^{\cdot} g_j d\eta = V_j, \tag{17a}$$

$$U_m \int_{-1}^{+1} f_m^{\cdot} g_j^{\cdot} d\eta + cU_{m,x}\int_{-1}^{+1} K_{\sigma\tau} f_m g_j^{\cdot} d\eta - cS_m \int_{-1}^{+1} K_{\epsilon\tau} g_m g_j^{\cdot} d\eta - Q_m \int_{-1}^{+1} K_{\gamma\tau} g_m^{\cdot} g_j^{\cdot} d\eta = -cV_{j,x}, \tag{17b}$$

$$cN_{xj,x} - Q_m \int_{-1}^{+1} g_m^{\cdot} f_j^{\cdot} d\eta = 0, \tag{17c}$$

$$Q_{j,x} + S_j = 0, \tag{17d}$$

where

$$N_{xj} = c\int_{-1}^{+1} f_j \frac{\partial W}{\partial u_{,x}} d\eta = cU_{m,x}\int_{-1}^{+1} K_{\sigma\epsilon} f_m f_j d\eta + cS_m \int_{-1}^{+1} K_{\sigma\sigma} g_m f_j d\eta + Q_m \int_{-1}^{+1} K_{\sigma\tau} g_m^{\cdot} f_j d\eta. \tag{17e}$$

The above holds for $j = 0, 1, 2, 3, \ldots$, with $S_0 = S_1 = Q_0 = V_0 = 0$ and $g_0 = g_1 = g_0^{\cdot} = 0$.

The associated boundary conditions will be, on the basis of (6b),

$$U_0 = U_1 = U_2 = \ldots = 0, \quad V_1 = V_2 = \ldots = 0, \quad \text{at} \quad x = 0, \tag{18a}$$

and, on the basis of (6c),

$$U_0 = \hat{U}_0, \quad U_1 = \hat{U}_1, \ldots, V_1 = \hat{V}_1 \ldots, \quad \text{at} \quad x = L. \tag{18b}$$

For (18b) to be valid it is necessary to have

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\hat{u}(y) = \hat{U}_0 f_0 + \hat{U}_1 f_1 + \ldots,$$

$$\hat{V}_j = \int_{-1}^{+1} \hat{v}(L, \eta) g_j^*(\eta) d\eta. \tag{19}$$

In connection with the problem as formulated we will be interested in particular in the associated end forces

$$\hat{N} = N_{x0}(L), \quad \hat{Q} = Q_1(L), \tag{20a}$$

and the associated end moment

$$\hat{M} = cN_{x1}(L). \tag{20b}$$

In what follows we consider the evaluation of the above system of equations for the three cases $J = 1, 2, 3$, in (11) and (12), subjected to the assumption of constant values of the constitutive coefficients K. The one-dimensional theory for the case with $J = 1$ will be designated as Timoshenko-type, or first-order theory.

Timoshenko-Type Beam Theory

The equilibrium equations of this theory follow from (17c,d) in the form

$$cN_{x0,x} = 0, \tag{21a}$$

$$cN_{x1,x} - Q_1 = 0, \tag{21b}$$

$$Q_{1,x} = 0. \tag{21c}$$

The constitutive relations follow from (17e) and (17b) in the form

$$N_{x0} = 2cK_{\sigma\epsilon}U_{0,x} + K_{\sigma\tau}Q_1, \tag{22a}$$

$$N_{x1} = \tfrac{2}{3} cK_{\sigma\epsilon}U_{1,x}, \tag{22b}$$

$$U_1 + cV_{1,x} + cK_{\sigma\tau}U_{0,x} = \tfrac{3}{5} K_{\gamma\tau}Q_1. \tag{22c}$$

The above may be recast in matrix form as

$$\begin{bmatrix} N_{x0} \\ N_{x1} \\ Q_1 \end{bmatrix} = \begin{bmatrix} C_{11}^T & 0 & C_{13}^T \\ 0 & C_{22}^T & 0 \\ C_{13}^T & 0 & C_{33}^T \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{1,x} \\ U_1 + cV_{1,x} \end{bmatrix}, \tag{23a}$$

where

$$C_{11}^T = 2E_{xx} + \frac{5}{3}\frac{E_{xx}^2}{G_x^2}\frac{G}{1 - E_{xx}G/G_x^2},$$

$$C_{13}^T = -\frac{5}{3}\frac{E_{xx}}{G_x}\frac{G}{1 - E_{xx}G/G_x^2},$$

$$C_{22}^T = \frac{2}{3}E_{xx}, \quad C_{33}^T = \frac{5}{3}\frac{G}{1 - E_{xx}G/G_x^2}. \tag{23b}$$

The corresponding boundary conditions are

$$U_0 = U_1 = V_1 = 0 \quad \text{at} \quad x = 0, \tag{24}$$

$$U_0 = \hat{U}_0, \quad U_1 = \hat{U}_1, \quad V_1 = \hat{V}_1 \quad \text{at} \quad x = L. \tag{25}$$

Remark. For an entirely displacement-based variational method, stiffer constitutive relations would have been obtained as shown in Appendix A. By comparing the constitutive relations (23) with (A8), one finds that the shear correction factor, $\frac{5}{6}$, appears not only in C_{33}^T as in the corresponding orthotropic problem but also in C_{11}^T and C_{13}^T.

The one-dimensional equations of the second-order theory and the third-order theory are summarized in Appendices B and C, respectively.

The boundary value problem defined by (21) and (23) – (25) is solved analytically as follows. By integrating (21a), (21c), and (21b), we obtain

$$N_{x0}(\xi) = \hat{N}, \tag{26a}$$

$$Q_1(\xi) = \hat{Q}, \tag{26b}$$

$$N_{x1}(\xi) = \frac{1}{c}\hat{M} - \hat{Q}\frac{1}{h}(1 - \xi), \tag{26c}$$

where $\xi = x/L$, and $h = c/L$. From (23) and (24) we find

$$U_0(\xi) = \frac{1}{2}\frac{1}{E_{xx}}\left(\hat{N}\frac{\xi}{h} + \frac{E_{xx}}{G_x}\hat{Q}\frac{\xi}{h}\right), \tag{27a}$$

$$U_1(\xi) = \frac{3}{2}\frac{1}{E_{xx}}\left\{\frac{1}{c}\hat{M}\frac{\xi}{h} - \hat{Q}\frac{1}{h^2}\left(\xi - \frac{\xi^2}{2}\right)\right\}, \tag{27b}$$

$$V_1(\xi) = \frac{1}{2}\frac{1}{E_{xx}}\left[\frac{E_{xx}}{G_x}\hat{N}\frac{\xi}{h} - \frac{3}{2}\frac{1}{c}\hat{M}\frac{\xi^2}{h^2} + \hat{Q}\left\{\left(\frac{6}{5}\frac{E_{xx}}{G} - \frac{1}{5}\frac{E_{xx}^2}{G_x^2}\right)\frac{\xi}{h} + \frac{3}{h^3}\left(\frac{\xi^2}{2} - \frac{\xi^3}{6}\right)\right\}\right]. \tag{27c}$$

With (25) the resulting deflection-force relationship can be written in the form

$$\begin{bmatrix} \hat{U}_0 \\ \hat{U}_1 \\ \hat{V}_1 \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & R_{13} \\ 0 & R_{22} & R_{23} \\ R_{13} & R_{23} & R_{33} \end{bmatrix} \begin{bmatrix} \hat{N} \\ \frac{1}{c}\hat{M} \\ \hat{Q} \end{bmatrix}, \tag{28a}$$

where

$$R_{11} = \frac{1}{2h}\frac{1}{E_{xx}}, \quad R_{13} = \frac{1}{2h}\frac{1}{G_x},$$

$$R_{22} = \frac{3}{2h}\frac{1}{E_{xx}}, \quad R_{23} = -\frac{3}{4h^2}\frac{1}{E_{xx}},$$

$$R_{33} = \frac{1}{2h^3}\frac{1}{E_{xx}} + \frac{3}{5h}\left(\frac{1}{G} - \frac{1}{6}\frac{E_{xx}}{G_x^2}\right). \tag{28b}$$

An inversion of (28a) gives the force-deflection relation

$$\begin{bmatrix} \hat{N} \\ \frac{1}{c}\hat{M} \\ \hat{Q} \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{12} & F_{22} & F_{23} \\ F_{13} & F_{23} & F_{33} \end{bmatrix} \begin{bmatrix} \hat{U}_0 \\ \hat{U}_1 \\ \hat{V}_1 \end{bmatrix}, \tag{29a}$$

where

$$F_{11} = 2hE_{xx} + \frac{8h^3}{\Delta}\frac{E_{xx}^3}{G_x^2}, \quad F_{12} = -\frac{4h^2}{\Delta}\frac{E_{xx}^2}{G_x},$$

$$F_{13} = -\frac{8h^3}{\Delta}\frac{E_{xx}^2}{G_x},$$

$$F_{22} = \frac{2h}{3}E_{xx}\left(1 + \frac{3}{\Delta}\right),$$

$$F_{23} = \frac{4h^2}{\Delta}E_{xx}, \quad F_{33} = \frac{8h^3}{\Delta}E_{xx}, \tag{29b}$$

and where

$$\Delta = 1 + \frac{24}{5}h^2\left(\frac{E_{xx}}{G} - \frac{E_{xx}^2}{G_x^2}\right). \tag{29c}$$

The boundary value problem defined by the second-order the-

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ory, (21), (B1), and (B4)–(B6) in Appendix B, has also been solved analytically with the help of a symbolic manipulator, Mathematica (Wolfram, 1991). The results are shown in Appendix D.

Compliances of Anisotropic Cantilever Beams

In order to demonstrate the capability of the proposed theories in capturing the effects of transverse shear and the normal-shear coupling, it is convenient to nondimensionalize compliances relative to those of the Euler-Bernoulli beam formulation, R^B_{ij}. The Bernoulli beam theory is based upon the displacement trial field with $\epsilon_y = \gamma = 0$, and the uniaxial stress state, $\sigma_x = K_{\sigma\epsilon}\epsilon_x$, with *constitutive* assumptions $\sigma_y = \tau = 0$. The compliance coefficients of the Bernoulli theory are

$$R^B_{11} = \frac{1}{2h}\frac{1}{E_{xx}}, \quad R^B_{13} = 0, \quad R^B_{22} = \frac{3}{2h}\frac{1}{E_{xx}},$$

$$R^B_{23} = -\frac{3}{4h^2}\frac{1}{E_{xx}}, \quad R^B_{33} = \frac{1}{2h^3}\frac{1}{E_{xx}}, \tag{30}$$

where $h = c/L$.

The nondimensionalized compliances of the first-order Timoshenko-type beam model, (28), become

$$\frac{R_{11}}{R^B_{11}} = \frac{R_{22}}{R^B_{22}} = \frac{R_{23}}{R^B_{23}} = 1, \quad \frac{R_{33}}{R^B_{33}} = 1 + h^2\left\{\frac{6}{5}\frac{E_{xx}}{G} - \frac{1}{5}\left(\frac{E_{xx}}{G_x}\right)^2\right\},$$

$$\frac{R_{13}}{R^B_{11}} = \frac{E_{xx}}{G_x}, \quad \frac{R_{13}}{R^B_{33}} = h^2\frac{E_{xx}}{G_x}. \tag{31}$$

In the above, E_{xx}/G represents the effect of transverse shear deformation, while E_{xx}/G_x represents the normal-shear stress coupling effect. For cantilever beams with anisotropic coupling, R_{33}/R^B_{33} in (31b) furnishes a prediction which is consistent with that furnished by Nair and Reissner (1976) for the limiting case of $E_{yy} \to \infty$.

Numerical Results

In order to examine the consequences of anisotropy with particular emphasis on the effect of constitutive coupling of normal and shear stresses, two materials, shown in Table 1, orthotropic with reference to axes x_1 and x_2 were selected. Material 1 is more strongly anisotropic than Material 2, as shown by the ratio of Young's moduli, E_{22}/E_{11}. By rotating the x_1-x_2 axes by an angle θ, a beam with axis in the x-direction becomes anisotropic with the desired constitutive coupling. The elastic moduli in (3) are computed by the following transformation relations:

$$\begin{bmatrix} \frac{1}{E_{xx}} & \frac{1}{E_{xy}} & \frac{1}{G_x} \\ \frac{1}{E_{xy}} & \frac{1}{E_{yy}} & \frac{1}{G_y} \\ \frac{1}{G_x} & \frac{1}{G_y} & \frac{1}{G} \end{bmatrix} = [O]^t \begin{bmatrix} \frac{1}{E_{11}} & \frac{-\nu_{12}}{E_{11}} & 0 \\ \frac{-\nu_{12}}{E_{11}} & \frac{1}{E_{22}} & 0 \\ 0 & 0 & \frac{1}{G_{12}} \end{bmatrix} [O], \tag{32a}$$

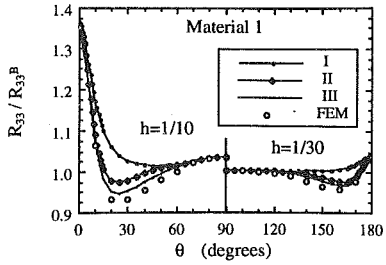


Fig. 1(a) R_{33}/R^B_{33} versus θ for Material 1 with $h = \frac{1}{10}$ and $\frac{1}{30}$

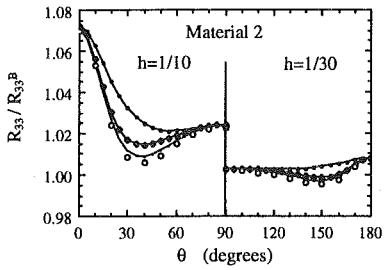


Fig. 1(b) R_{33}/R^B_{33} versus θ for Material 2 with $h = \frac{1}{10}$ and $\frac{1}{30}$

where

$$[O] = \begin{bmatrix} \frac{1}{2}(1 + \cos 2\theta) & \frac{1}{2}(1 - \cos 2\theta) & \sin 2\theta \\ \frac{1}{2}(1 - \cos 2\theta) & \frac{1}{2}(1 + \cos 2\theta) & -\sin 2\theta \\ -\frac{1}{2}\sin 2\theta & \frac{1}{2}\sin 2\theta & \cos 2\theta \end{bmatrix}, \tag{32b}$$

and where $[\ ]^t$ denotes the transpose of the matrix $[\ ]$. When $\theta = 0$ deg the strong fiber direction coincides with the x-axis, and when $\theta = 90$ deg it coincides with the y-axis.

Compliances have been computed for the first-order, second-order, and third-order beam theories. These results are shown in figures with the numbers I, II, and III. It is noted that the compliance curves are symmetric with respect to $\theta = 90$ deg. Comparisons are made for diagonal compliance terms, R_{11}, R_{22}, and R_{33}, as well as for the off-diagonal terms R_{13} and R_{23}. These compliances are nondimensionalized with respect to those in (30). However, R_{13} is nondimensionalized by R_{33} since $R^B_{13} = 0$. The comparisons of compliances were made only for the three by three elements, (28), associated with the first-order theory. However, due to higher-order displacements and moments, the compliance matrices of the second and the third-order theories are, respectively, five by five and seven by seven. Therefore, the compliances of the second and the third-order theories may be compared with those of the first-order theory by setting those higher-order moments to zero.

Figures 1(a) and (b) illustrate the variation of the compliance R_{33}/R^B_{33} due to a transverse load $\hat{Q}$ as a function of the rotation angle θ, for Materials 1 and 2, respectively. The results for a thick beam with $h = \frac{1}{10}$ are shown for $\theta = 0$ deg to 90 deg, and those for a thin beam with $h = \frac{1}{30}$ are shown for $\theta = 90$ deg to 180 deg. The closeness of the predictions by the three shear deformable beam models, I, II, and III, demonstrates the accuracy of the first-order beam model I. Finite element calculation was conducted for Material 1 and 2 to verify the accuracy

Table 1 Material properties

Material	Material 1	Material 2
E_{22}/E_{11}	0.1	0.3333
G_{12}/E_{11}	0.03333	0.1667
ν_{12}	0.3	0.25

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of model III. A general-purpose finite element code, MARC (1992), was employed for the numerical analyses. The rectangular domain was discretized by using square, eight-node quadrilateral elements. Meshes included eight layers for $h = \frac{1}{10}$ and four layers for $h = \frac{1}{30}$ of square elements in the thickness direction. In order to reflect the vanishing higher-order moments of the second and the third-order theories, the following loading conditions at $x = L$ were employed for the calculation of R_{33}:

$$\sigma_x = 0, \quad v = 1 \quad \text{at} \quad x = L. \tag{33}$$

By comparing the compliance curves in the figures for a thick beam, $h = \frac{1}{10}$, and for a thin beam, $h = \frac{1}{30}$, the transverse shear effect diminishes as the thickness to span ratio $2h$ $(=2c/L)$ reduces. Further, the transverse shear effect becomes maximum near $\theta = 0$ deg. By comparing the compliances at $\theta = 0$ deg between Figs. 1(a) and (b), it can be seen that a stronger anisotropy yields a stronger effect of transverse shear deformation. Since Material 1 has $E_{22}/E_{11} = 0.1$, the results of Fig. 1(a) near $\theta = 0$ deg and 90 deg mimic the cases, respectively, for $E_{xx} \gg E_{yy}$ and $E_{xx} \ll E_{yy}$. At these limiting cases, the results of Fig. 1(a) confirm the conclusion drawn by Nair and Reissner (1975) who reported a remarkable sensitivity of the compliance when $E_{xx} \gg E_{yy}$ and an insensitivity when $E_{xx} \ll E_{yy}$. In addition, for the limiting case of $E_{xx} \ll E_{yy}$ Nair and Reissner (1976) reported that the normal and shear stress coupling reduces the compliance quadratically. This prediction is also confirmed near $\theta = 90$ deg since the coupling increases as θ deviates from 90 deg in Figs. 1. It is further noted that the predictions by the first-order theory, I, considerably deviates from those of higher-order theories, II and III, for stiffer (Material 1) and thicker ($h = \frac{1}{10}$) beams at the angles where the compliance takes on minimum values.

Figures 2(a) and (b) show, for Materials 1 and 2, respectively, the variation of R_{13}/R_{33} with θ for both a thick beam and a thin beam. This term, R_{13}/R_{33}, represents the normal-shear coupling effect. The Bernoulli beam theory fails to capture this coupling effect, while all three beam theories predict significant coupling effects near $\theta = 10$ deg (Material 1) and 20 deg (Material 2). At these angles the normal-shear coupling term E_{xx}/G_x takes maximum values. The absolute value of the coupling compliance increases with increasing $h = c/L$.

Figure 3 illustrates the variation of R_{22}/R_{22}^B with θ for Material 1. For Material 2 the deviation of R_{22}/R_{22}^B from unity becomes almost negligible. The variation of the compliance curves II and III with θ is similar to that of Figs. 2 and is characteristic for the normal-shear coupling. The first-order beam model I fails to capture this coupling effect. However, it is seen that the coupling effect becomes only noticeable for materials of extremely strong anisotropy. Figure 4 shows the variation of R_{23}/R_{23}^B with θ for Material 1. The same observation as Fig. 3 is drawn from Fig. 4. The beam models II and III predict the minor normal-shear coupling effect. For Material 2 R_{23}/R_{23}^B remains almost unity.

The variation of R_{11}/R_{11}^B with θ for both $h = \frac{1}{10}$ and $\frac{1}{30}$ does not demonstrate noticeable deviation from unity. In Figs. 3 and 4, it is interesting to note that the higher-order beam models II and III can capture even minute normal-shear coupling effects.

In summary, the numerical results show that the normal-shear coupling effect becomes important in R_{33}, R_{22}, and R_{13}.

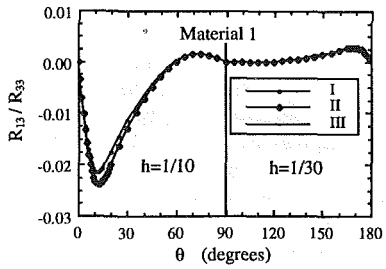


Fig. 2(a) R_{13}/R_{33} versus θ for Material 1 with $h = \frac{1}{10}$ and $\frac{1}{30}$

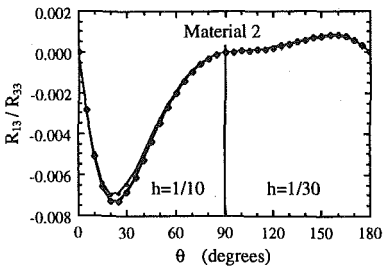


Fig. 2(b) R_{13}/R_{33} versus θ for Material 2 with $h = \frac{1}{10}$ and $\frac{1}{30}$

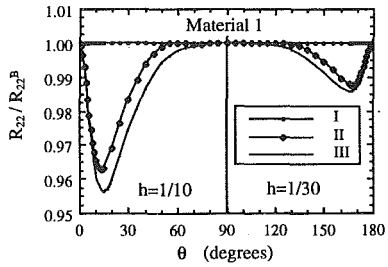


Fig. 3 R_{22}/R_{22}^B versus θ for Material 1 with $h = \frac{1}{10}$ and $\frac{1}{30}$

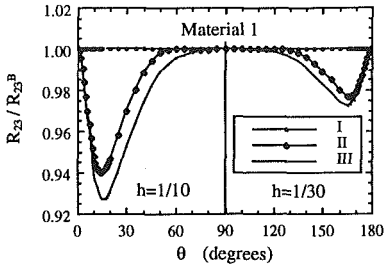


Fig. 4 R_{23}/R_{23}^B versus θ for Material 1 with $h = \frac{1}{10}$ and $\frac{1}{30}$

References

Hashin, Z., 1967, "Plane Anisotropic Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 257–262.

MARC Analysis Research Corporation, 1992, *MARC Reference Library, Volume A: User Information,* Rev. K.5, Palo Alto, CA.

Hildebrand, F. B., 1943, "On the Stress Distribution in Cantilever Beams," *Journal of Mathematics and Physics,* Vol. 22, pp. 188–203.

Nair, S., and Reissner, E., 1975, "Improved Upper and Lower Bounds for Deflections of Orthotropic Cantilever Beams," *International Journal of Solids and Structures,* Vol. 11, pp. 961–971.

Nair, S., and Reissner, E., 1976, "On the Determination of Stresses and Deflections for Anisotropic Homogeneous Cantilever Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 75–80.

Reissner, E., 1940, "A Contribution to the Theory of Elasticity of Non-Isotropic Materials (with Applications to Problems of Bending and Torsion)," *Philosophical Magazine and Journal of Science,* Ser. 7, Vol. 30, pp. 418–427.

Reissner, E., 1984, "On a Certain Mixed Variational Theorem and a Proposed Application," *International Journal for Numerical Methods in Engineering,* Vol. 20, pp. 1366–1368.

Wolfram, S., 1991, *Mathematica: A System for Doing Mathematics by Computer,* 2nd ed., Addison-Wesley, Redwood City, CA.

APPENDIX A

Use of the Minimum Potential Energy Theorem

The trial displacement field

$$u(x, \eta) \approx U_0(x) + U_1(x)\eta,$$
$$v(x, \eta) \approx V_1(x), \tag{A1}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

is employed together with the *constitutive* assumption $\sigma_y = 0$. With which (3) becomes

$$\begin{bmatrix} \sigma_x \\ \tau \end{bmatrix} = \begin{bmatrix} E_{xx} + \frac{E_{xx}^2}{G_x^2}\hat{G} & -\frac{E_{xx}}{G_x}\hat{G} \\ -\frac{E_{xx}}{G_x}\hat{G} & \hat{G} \end{bmatrix} \begin{bmatrix} \epsilon_x \\ \gamma \end{bmatrix}, \quad \text{(A2}a\text{)}$$

where

$$\hat{G} = \frac{G}{1 - E_{xx}G/G_x^2}. \quad \text{(A2}b\text{)}$$

The strains are computed from (A1) and (2) as

$$\epsilon_x = U_{0,x} + U_{1,x}\eta, \quad \gamma = V_{1,x} + \frac{1}{c}U_1. \quad \text{(A3)}$$

The strain energy function $W_d(\epsilon_x, \gamma)$ is defined as

$$W_d = \frac{1}{2}[\epsilon_x \quad \gamma] \begin{bmatrix} E_{xx} + \frac{E_{xx}^2}{G_x^2}\hat{G} & -\frac{E_{xx}}{G_x}\hat{G} \\ -\frac{E_{xx}}{G_x}\hat{G} & \hat{G} \end{bmatrix} \begin{bmatrix} \epsilon_x \\ \gamma \end{bmatrix}. \quad \text{(A4)}$$

The minimum potential energy theorem furnishes

$$\iint \left\{ \frac{\partial W_d}{\partial \epsilon_x}\delta u_{,x} + \frac{\partial W_d}{\partial \gamma}(\delta u_{,y} + \delta v_{,x}) \right\} dxdy = 0, \quad \text{(A5)}$$

where

$$\sigma_x = \frac{\partial W_d}{\partial \epsilon_x}, \quad \tau = \frac{\partial W_d}{\partial \gamma}. \quad \text{(A6)}$$

The substitution of (A1) and (A3) into (A5) yields (21a,b,c) as the Euler equations together with

$$(N_{x0}, N_{x1}) = c\int_{-1}^{+1} \frac{\partial W_d}{\partial \epsilon_x}(1, \eta)d\eta, \quad \text{(A7}a, b\text{)}$$

and

$$Q_1 = c\int_{-1}^{+1} \frac{\partial W_d}{\partial \gamma}d\eta. \quad \text{(A7}c\text{)}$$

The substitution of (A4) and (A3) into (A7) furnishes the constitutive relations

$$\begin{bmatrix} N_{x0} \\ N_{x1} \\ Q_1 \end{bmatrix} = \begin{bmatrix} 2E_{xx} + 2\frac{E_{xx}^2}{G_x^2}\hat{G} & 0 & -2\frac{E_{xx}}{G_x}\hat{G} \\ 0 & \frac{2}{3}E_{xx} & 0 \\ -2\frac{E_{xx}}{G_x}\hat{G} & 0 & 2\hat{G} \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{1,x} \\ U_1 + cV_{1,x} \end{bmatrix}, \quad \text{(A8)}$$

which is to be compared to (23a,b), taking into account of (A2b).

APPENDIX B

The Second-Order Theory

In accordance with (17c,d), the equilibrium equations of this theory are (21a,b,c), complemented by

$$cN_{x2,x} - \tfrac{4}{5}Q_2 = 0, \quad Q_{2,x} + S_2 = 0. \quad \text{(B1}a, b\text{)}$$

The associated constitutive relations follow from (17e) and (17a,b) in the form

$$\begin{bmatrix} N_{x0} \\ N_{x1} \\ N_{x2} \end{bmatrix} = K_{\sigma\epsilon} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{1,x} \\ cU_{2,x} \end{bmatrix} + \begin{bmatrix} K_{\sigma\tau} & 0 & -\frac{4}{15}K_{\sigma\sigma} \\ 0 & \frac{4}{15}K_{\sigma\tau} & 0 \\ -\frac{1}{5}K_{\sigma\tau} & 0 & \frac{8}{105}K_{\sigma\sigma} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ cS_2 \end{bmatrix}, \quad \text{(B2)}$$

and

$$\begin{bmatrix} \frac{3}{5}K_{\gamma\tau} & -\frac{6}{35}K_{\epsilon\tau} \\ -\frac{6}{35}K_{\epsilon\tau} & \frac{16}{315}K_{\epsilon\sigma} \end{bmatrix} \begin{bmatrix} Q_1 \\ cS_2 \end{bmatrix} = \begin{bmatrix} K_{\sigma\tau} & -\frac{1}{5}K_{\sigma\tau} & 1 & 0 \\ -\frac{4}{15}K_{\sigma\sigma} & \frac{8}{105}K_{\sigma\sigma} & 0 & 1 \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{2,x} \\ U_1 + cV_{1,x} \\ -V_2 \end{bmatrix}, \quad \text{(B3}a\text{)}$$

$$Q_2 = \frac{105}{16K_{\gamma\tau}} \begin{bmatrix} \frac{4}{15}K_{\sigma\tau} & 1 \end{bmatrix} \begin{bmatrix} cU_{1,x} \\ \frac{4}{5}U_2 + cV_{2,x} \end{bmatrix}. \quad \text{(B3}b\text{)}$$

Equations (B2) and (B3) may be combined to yield the constitutive relations for spanwise stress and displacement-variables as follows:

$$\{N\} = [C]\{e\}, \quad \text{(B4}a\text{)}$$

where

$$\{N\}^t = [N_{x0} \quad N_{x1} \quad N_{x2} \quad Q_1 \quad Q_2 \quad cS_2], \quad \text{(B4}b\text{)}$$

$$\{e\}^t = [cU_{0,x} \quad cU_{1,x} \quad cU_{2,x} \quad U_1 + cV_{1,x} \quad \tfrac{4}{5}U_2 + cV_{2,x} \quad -V_2], \quad \text{(B4}c\text{)}$$

and

$$[C] = \begin{bmatrix} C_{11} & 0 & C_{13} & C_{14} & 0 & C_{16} \\ 0 & C_{22} & 0 & 0 & C_{25} & 0 \\ C_{13} & 0 & C_{33} & C_{34} & 0 & C_{36} \\ C_{14} & 0 & C_{34} & C_{44} & 0 & C_{46} \\ 0 & C_{25} & 0 & 0 & C_{55} & 0 \\ C_{16} & 0 & C_{36} & C_{46} & 0 & C_{66} \end{bmatrix}. \quad \text{(B4}d\text{)}$$

The nonzero elements C_{ij} in this are

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$C_{11} = 2K_{\sigma\epsilon} + C_{inv}\left(\frac{5}{3}K_{\sigma\tau}^2K_{\epsilon\sigma} - 3K_{\sigma\tau}K_{\sigma\sigma}K_{\epsilon\tau} + \frac{7}{5}K_{\sigma\sigma}^2K_{\gamma\tau}\right),$$

$$C_{13} = C_{inv}\left(-\frac{1}{3}K_{\sigma\tau}^2K_{\epsilon\sigma} + \frac{51}{70}K_{\sigma\tau}K_{\sigma\sigma}K_{\epsilon\tau} - \frac{2}{5}K_{\sigma\sigma}^2K_{\gamma\tau}\right),$$

$$C_{14} = C_{inv}\left(\frac{5}{3}K_{\sigma\tau}K_{\epsilon\sigma} - \frac{3}{2}K_{\sigma\sigma}K_{\epsilon\tau}\right),$$

$$C_{16} = C_{inv}\left(\frac{45}{8}K_{\sigma\tau}K_{\epsilon\tau} - \frac{21}{4}K_{\sigma\sigma}K_{\gamma\tau}\right),$$

$$C_{22} = \frac{2}{3}K_{\sigma\epsilon} + \frac{7}{15}\frac{K_{\sigma\tau}^2}{K_{\gamma\tau}}, \quad C_{25} = \frac{7}{4}\frac{K_{\sigma\tau}}{K_{\gamma\tau}},$$

$$C_{33} = \frac{2}{5}K_{\sigma\epsilon} + C_{inv}\left(\frac{1}{15}K_{\sigma\tau}^2K_{\epsilon\sigma} + \frac{4}{35}K_{\sigma\sigma}^2K_{\gamma\tau} - \frac{6}{35}K_{\sigma\tau}K_{\sigma\sigma}K_{\epsilon\tau}\right),$$

$$C_{34} = C_{inv}\left(-\frac{1}{3}K_{\sigma\tau}K_{\epsilon\sigma} + \frac{3}{7}K_{\sigma\sigma}K_{\epsilon\tau}\right),$$

$$C_{36} = C_{inv}\left(-\frac{9}{8}K_{\sigma\tau}K_{\epsilon\tau} + \frac{3}{2}K_{\sigma\sigma}K_{\gamma\tau}\right), \quad C_{44} = C_{inv}\frac{5}{3}K_{\epsilon\sigma},$$

$$C_{46} = C_{inv}\frac{45}{8}K_{\epsilon\tau}, \quad C_{55} = \frac{105}{16}\frac{1}{K_{\gamma\tau}},$$

$$C_{66} = C_{inv}\frac{315}{16}K_{\gamma\tau}, \qquad \text{(B4}e\text{)}$$

where

$$C_{inv} = \frac{1}{K_{\gamma\tau}K_{\epsilon\tau} - \frac{27}{28}K_{\epsilon\tau}^2}. \qquad \text{(B4}f\text{)}$$

The above result was also checked with Mathematica (Wolfram, 1991).

The corresponding boundary conditions are

$$U_0 = U_1 = U_2 = V_1 = V_2 = 0 \quad \text{at} \quad x = 0, \qquad \text{(B5)}$$

and

$$U_0 = \hat{U}_0, \quad U_1 = \hat{U}_1, \quad U_2 = \hat{U}_2,$$
$$V_1 = \hat{V}_1, \quad V_2 = \hat{V}_2 \quad \text{at} \quad x = L. \qquad \text{(B6)}$$

It is noted here that for orthotropic materials, $K_{\sigma\tau} = K_{\epsilon\tau} = 0$ in (4), Eq. (B4d) simplifies as $C_{14} = C_{25} = C_{34} = C_{46} = 0$.

APPENDIX C

The Third-Order Theory

The equilibrium equations consist of (21a,b,c), (B1a,b) and

$$cN_{x3,x} + \frac{16}{7}Q_3 = 0, \quad Q_{3,x} + S_3 = 0. \qquad \text{(C1}a, b\text{)}$$

The associated constitutive relations follow from (17a,b) and (17e) in the form

$$\begin{bmatrix} \frac{3}{5}K_{\gamma\tau} & -\frac{6}{35}K_{\epsilon\tau} & \frac{8}{35}K_{\gamma\tau} \\ -\frac{6}{35}K_{\epsilon\tau} & \frac{16}{315}K_{\epsilon\sigma} & -\frac{32}{315}K_{\epsilon\tau} \\ \frac{8}{35}K_{\gamma\tau} & -\frac{32}{315}K_{\epsilon\tau} & \frac{256}{315}K_{\gamma\tau} \end{bmatrix} \begin{bmatrix} Q_1 \\ cS_2 \\ Q_3 \end{bmatrix} = \begin{bmatrix} K_{\sigma\tau} & -\frac{1}{5}K_{\sigma\tau} & 1 & 0 & 0 \\ -\frac{4}{15}K_{\sigma\sigma} & \frac{8}{105}K_{\sigma\sigma} & 0 & 1 & 0 \\ 0 & -\frac{16}{35}K_{\sigma\tau} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{2,x} \\ U_1 + cV_{1,x} \\ -V_2 \\ -\frac{16}{7}U_3 + cV_{3,x} \end{bmatrix}, \qquad \text{(C2}a\text{)}$$

$$\begin{bmatrix} \frac{16}{105}K_{\gamma\tau} & \frac{32}{315}K_{\epsilon\tau} \\ \frac{32}{315}K_{\epsilon\tau} & \frac{256}{3465}K_{\epsilon\sigma} \end{bmatrix} \begin{bmatrix} Q_2 \\ cS_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{15}K_{\sigma\tau} & -\frac{4}{35}K_{\sigma\tau} & 1 & 0 \\ \frac{16}{105}K_{\sigma\sigma} & -\frac{32}{315}K_{\sigma\sigma} & 0 & 1 \end{bmatrix} \begin{bmatrix} cU_{1,x} \\ cU_{3,x} \\ \frac{4}{5}U_2 + cV_{2,x} \\ -V_3 \end{bmatrix}, \qquad \text{(C2}b\text{)}$$

and

$$\begin{bmatrix} N_{x0} \\ N_{x1} \\ N_{x2} \\ N_{x3} \end{bmatrix} = K_{\sigma\epsilon}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{2}{5} & 0 \\ 0 & 0 & 0 & \frac{2}{7} \end{bmatrix} \begin{bmatrix} cU_{0,x} \\ cU_{1,x} \\ cU_{2,x} \\ cU_{3,x} \end{bmatrix} + \begin{bmatrix} K_{\sigma\tau} & 0 & -\frac{4}{15}K_{\sigma\sigma} & 0 & 0 \\ 0 & \frac{4}{15}K_{\sigma\tau} & 0 & 0 & \frac{16}{105}K_{\sigma\sigma} \\ -\frac{1}{5}K_{\sigma\tau} & 0 & \frac{8}{105}K_{\sigma\sigma} & -\frac{16}{35}K_{\sigma\tau} & 0 \\ 0 & -\frac{4}{35}K_{\sigma\tau} & 0 & 0 & -\frac{32}{315}K_{\sigma\sigma} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ cS_2 \\ Q_3 \\ cS_3 \end{bmatrix}. \qquad \text{(C3)}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The corresponding boundary conditions are

$$U_0 = U_1 = U_2 = U_3 = V_1 = V_2 = V_3 = 0 \quad \text{at} \quad x = 0, \quad \text{(C4)}$$

and

$$U_0 = \hat{U}_0, \quad U_1 = \hat{U}_1, \quad U_2 = \hat{U}_2, \quad U_3 = \hat{U}_3,$$
$$V_1 = \hat{V}_1, \quad V_2 = \hat{V}_2, \quad V_3 = \hat{V}_3 \quad \text{at} \quad x = L. \quad \text{(C5)}$$

APPENDIX D

The Boundary Value Problem of the Second-Order Theory

This problem becomes statically indeterminate, since we cannot integrate (B1a) and (B1b). We can integrate (21a), (21b), and (21c) to find (26). In order to integrate (B1a) and (B1b), we first use (26) to eliminate $hU_{0,\xi}$, $hU_{1,\xi}$, and $U_1 + hV_{1,\xi}$ in terms of the constants $\hat{N}$, $\hat{M}/c$, $\hat{Q}$, and the other strain components as follows:

$$hU_{1,\xi} = \frac{1}{C_{22}}\left\{-\hat{Q}\frac{1}{h}(1-\xi) + \frac{1}{c}\hat{M} - C_{25}\left(\frac{4}{5}U_2 + hV_{2,\xi}\right)\right\}, \quad \text{(D1)}$$

$$\begin{bmatrix} hU_{0,\xi} \\ U_1 + hV_{1,\xi} \end{bmatrix} = \frac{\hat{N}}{C_m^2}\begin{bmatrix} C_{44} \\ -C_{14} \end{bmatrix} + \frac{\hat{Q}}{C_m^2}\begin{bmatrix} -C_{14} \\ C_{11} \end{bmatrix} - \frac{1}{C_m^2}\begin{bmatrix} C_{44} & -C_{14} \\ -C_{14} & C_{11} \end{bmatrix}\begin{bmatrix} C_{13} & C_{16} \\ C_{34} & C_{46} \end{bmatrix}\begin{bmatrix} hU_{2,\xi} \\ -V_2 \end{bmatrix}, \quad \text{(D2)}$$

where

$$C_m^2 = C_{11}C_{44} - C_{14}^2. \quad \text{(D3)}$$

We then rewrite the constitutive relations for N_{x2}, Q_2, and S_2 of (B4a) by incorporating (D2) as

$$Q_2 = \hat{C}_{55}\left(\frac{4}{5}U_2 + hV_{2,\xi}\right) - \hat{Q}\frac{C_{25}}{C_{22}}\frac{1}{h}(1-\xi) + \frac{C_{25}}{C_{22}}\frac{1}{c}\hat{M}, \quad \text{(D4)}$$

$$\begin{bmatrix} N_{x2} \\ cS_2 \end{bmatrix} = \begin{bmatrix} \hat{C}_{33} & \hat{C}_{36} \\ \hat{C}_{36} & \hat{C}_{66} \end{bmatrix}\begin{bmatrix} hU_{2,\xi} \\ -V_2 \end{bmatrix} + \frac{\hat{N}}{C_m^2}\begin{bmatrix} C_{13} & C_{34} \\ C_{16} & C_{46} \end{bmatrix}\begin{bmatrix} C_{44} \\ -C_{14} \end{bmatrix} + \frac{\hat{Q}}{C_m^2}\begin{bmatrix} C_{13} & C_{34} \\ C_{16} & C_{46} \end{bmatrix}\begin{bmatrix} -C_{14} \\ C_{11} \end{bmatrix}, \quad \text{(D5)}$$

where

$$\begin{bmatrix} \hat{C}_{33} & \hat{C}_{36} \\ \hat{C}_{36} & \hat{C}_{66} \end{bmatrix} = \begin{bmatrix} C_{33} & C_{36} \\ C_{36} & C_{66} \end{bmatrix} - \frac{1}{C_m^2}\begin{bmatrix} C_{13} & C_{34} \\ C_{16} & C_{46} \end{bmatrix}\begin{bmatrix} C_{44} & -C_{14} \\ -C_{14} & C_{11} \end{bmatrix}\begin{bmatrix} C_{13} & C_{16} \\ C_{34} & C_{46} \end{bmatrix}, \quad \text{(D6)}$$

$$\hat{C}_{55} = C_{55} - \frac{C_{25}^2}{C_{22}}. \quad \text{(D7)}$$

By substituting (D4) and (D5) into the remaining equilibrium equations, (B1a) and (B1b), we obtain the following system of ordinary differential equations for U_2 and V_2:

$$\begin{bmatrix} \hat{C}_{33}h^2\frac{d^2}{d\xi^2} - \frac{16}{25}\hat{C}_{55} & -\left(\hat{C}_{36} + \frac{4}{5}\hat{C}_{55}\right)h\frac{d}{d\xi} \\ -\left(\hat{C}_{36} + \frac{4}{5}\hat{C}_{55}\right)h\frac{d}{d\xi} & -\hat{C}_{55}h^2\frac{d^2}{d\xi^2} + \hat{C}_{66} \end{bmatrix}\begin{bmatrix} U_2 \\ V_2 \end{bmatrix}$$
$$= \hat{N}\begin{bmatrix} 0 \\ \frac{1}{C_m^2}(C_{16}C_{44} - C_{46}C_{14}) \end{bmatrix} + \hat{Q}\begin{bmatrix} -\frac{4}{5}\frac{C_{25}}{C_{22}}\frac{1}{h}(1-\xi) \\ \frac{C_{25}}{C_{22}} + \frac{1}{C_m^2}(C_{11}C_{46} - C_{14}C_{16}) \end{bmatrix} + \frac{1}{c}\hat{M}\begin{bmatrix} \frac{4}{5}\frac{C_{25}}{C_{22}} \\ 0 \end{bmatrix}. \quad \text{(D8)}$$

In order to find homogeneous solutions of (D8), we try

$$\begin{bmatrix} U_2 \\ V_2 \end{bmatrix}^h = \begin{bmatrix} A \\ B \end{bmatrix}e^{s\xi}, \quad \text{(D9)}$$

and find the characteristic equation for s

$$\hat{C}_{33}\hat{C}_{55}(hs)^4 + \left\{\left(\hat{C}_{36} + \frac{4}{5}\hat{C}_{55}\right)^2 - \hat{C}_{33}\hat{C}_{66} - \frac{16}{25}\hat{C}_{55}^2\right\}(hs)^2 + \frac{16}{25}\hat{C}_{55}\hat{C}_{66} = 0. \quad \text{(D10)}$$

Since the above coefficients of the quadratic equation for $(hs)^2$ are real, we obtain a complex conjugate pair, say, $\hat{s}^2$ and $(\hat{s}^2)^*$. By taking the square roots, we find that zeros of (D10) are $\hat{s}$, $-\hat{s}$, $\hat{s}^*$ and $-\hat{s}^*$. The real part of $\hat{s}$ represents the boundary layer effect at $\xi = 0$ or 1.

The general solution of (D8) can be written as

$$\begin{bmatrix} U_2 \\ V_2 \end{bmatrix} = \alpha^{(1)}\begin{bmatrix} A^{(1)} \\ B^{(1)} \end{bmatrix}e^{-\hat{s}(1-\xi)} + \alpha^{(1)*}\begin{bmatrix} A^{(1)} \\ B^{(1)} \end{bmatrix}^* e^{-\hat{s}^*(1-\xi)} + \alpha^{(2)}\begin{bmatrix} A^{(2)} \\ B^{(2)} \end{bmatrix}e^{-\hat{s}\xi} + \alpha^{(2)*}\begin{bmatrix} A^{(2)} \\ B^{(2)} \end{bmatrix}^* e^{-\hat{s}^*\xi}$$
$$+ \hat{N}\begin{bmatrix} 0 \\ \frac{1}{\hat{C}_{66}}\frac{1}{C_m^2}(C_{16}C_{44} - C_{14}C_{46}) \end{bmatrix} + \frac{1}{c}\hat{M}\begin{bmatrix} -\frac{5}{4\hat{C}_{55}}\frac{C_{25}}{C_{22}} \\ 0 \end{bmatrix}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+ \hat{Q} \begin{bmatrix} \dfrac{5}{4\hat{C}_{55}} \dfrac{C_{25}}{C_{22}} \dfrac{1}{h} (1 - \xi) \\ \dfrac{1}{\hat{C}_{66}} \dfrac{1}{C_m^2} (C_{11}C_{46} - C_{14}C_{16}) - \dfrac{5\hat{C}_{36}}{4\hat{C}_{66}\hat{C}_{55}} \dfrac{C_{25}}{C_{22}} \end{bmatrix}, \quad \text{(D11)}$$

where $\alpha^{(1)}$ and $\alpha^{(2)}$ are the amplitudes of the mode shapes $[A^{(1)} \quad B^{(1)}]^T$ and $[A^{(2)} \quad B^{(2)}]^T$, respectively.

The $\hat{N}$-, $\hat{M}/c$-, and $\hat{Q}$-terms of (D11) represent particular solutions of the corresponding term of (D8).

The real and imaginary parts of $\alpha^{(1)}$ and $\alpha^{(2)}$ are determined by the remaining boundary conditions:

$$U_2 = V_2 = 0 \quad \text{at} \quad \xi = 0. \quad \text{(D12)}$$

With the solution for $U_2(\xi)$ and $V_2(\xi)$ one can integrate (D1) and (D2) for U_0, U_1, and V_1 with

$$U_0 = U_1 = V_1 = 0 \quad \text{at} \quad \xi = 0. \quad \text{(D13)}$$

From the general solution for displacements one can obtain a compliance relation similar to (28).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

H. P. Gavin
Assistant Professor,
Department of Civil and
Environmental Engineering,
Duke University,
Durham, NC 27708-0287
Assoc. Mem. ASME

R. D. Hanson
FEMA Disaster Field Office,
245 S. Los Robles, Room 630,
Pasadena, CA 91102-6020

F. E. Filisko
Professor,
Department of Materials Science
and Engineering,
The University of Michigan,
Ann Arbor, MI 48109

Electrorheological Dampers, Part I: Analysis and Design

Electrorheological (ER) materials are suspensions of specialized, micron-sized particles in nonconducting oils. When electric fields are applied to ER materials, they exhibit dramatic changes (within milli-seconds) in material properties. Pre-yield, yielding, and post-yield mechanisms are all influenced by the electric field. Namely, an applied electric field dramatically increases the stiffness and energy dissipation properties of these materials. A previously known cubic equation which describes the flow of fluids with a yield stress through a rectangular duct can be applied to annular flow, provided that certain conditions on the material properties are satisfied. An analytic solution and a uniform approximation to the solution, for the rectangular duct Poiseuille flow case is presented. A numerical method is required to solve the flow in annular geometries. The approximation for rectangular ducts is extended to deal with the annular duct case.

1 Introduction

Electrorheological (ER) materials are suspensions of unique micron-sized particles in hydrophobic, insulating, dielectric oils. The composition of ER materials can vary widely (Block and Kelly, 1988; Block et al., 1990; Jordan and Shaw, 1989; Weiss et al., 1993), but almost all materials make use of particles which become macroscopic dipoles in the presence of strong electric fields (kV/mm). ER materials exhibit order-of-magnitude and reversible changes in behavior when subjected to strong electric fields. Typically, ER material behavior changes within 1–10 milli-seconds upon application or removal of the electric field.

Interactions among the dispersed particles, fluid flow, mobile ions, and electrostatic fields contribute to the changes in material properties. When an electric field is applied to an ER suspension, the particles polarize and align themselves in the direction of the electric field. On a longer time-scale, these fibrils drift together and bridge the electrode gap. As the electric field through an ER material increases, the yield stress, stiffness, and viscous damping increase, while the yield strain decreases. Materials with properties which can change as a function of an electrical stimulus are attractive candidates for use in vibration control devices.

Many theories of ER activity are based on idealized electrostatic interactions. Particles aligned along the electric field are attracted to each other; those arranged across the field are repelled (Halsey and Martin, 1993). Theoretically, the polarization force is long-range, quadratically related to the field, and is very sensitive to the particle size (Gavin, 1994). According to point-dipole pair interaction energy models, ER mechanical properties (i.e., stiffness and yield stress) should be proportional to the square of the electric field strength. The exponents of power laws relating yield stresses to electric fields in actual ER suspensions range from 1.2 to 2.5. By examining the derivation of the two major components (particle polarization and particle interaction) of simplified electrostatic theories of electrorheology, inappropriate assumptions have been identified (Gavin, 1994; Klingenberg, 1993). Regarding particle polarization: in concentrated ER suspensions, the field near the particle surface is concentrated by a factor of 10 to 20, due to the presence of neighboring polarized particles (Chen, Sprecher, and Conrad, 1991). If water, salts, or other ionic impurities exist in the dispersant, the field-induced dipole formed by interfacial polarization will almost completely mask the dipole formed by bulk polarization of the particle (Filisko, 1991). The dielectric constant of the particle, in this case, is almost meaningless. Also, many ER materials are made of nonspherical convoluted particles. Regarding particle interactions: many simplified electrostatic theories assume that the particles are widely spaced, and aligned with a uniform electric field. These assumptions are not warranted in most ER suspensions.

More detailed models for describing ER phenomena have focused on improving the assumption on particle spacing (Klingenberg et al., 1991), the electric field configuration (Bonnecaze, 1992; Chen, Sprecher, and Conrad, 1991; Klingenberg et al., 1993), and polarization mechanisms (Davis, 1993). Some of these kinematic models have successfully described the behavior of certain ER materials (Brooks et al., 1986; Marshall et al., 1989). Studies related to the effect of dielectric loss, the effect of mobile ions, or how the morphology influences polarization and the ER effect address the current research needs relating the these materials (Conrad and Sprecher, 1991; Filisko and Radzilowsky, 1990; Filisko, 1991; Filisko and Wang, 1994; Schubring and Filisko, 1994).

Primitive ER suspensions can be fabricated by mixing starch, cellulose, or silica, with paraffin oil. The hydrophilic nature of corn starch and the polar nature of water molecules is a likely source of the strong ER activity sometimes observed in such materials. Water adsorbed into the dispersed phase of these ER materials allows electrical currents to flow between the electrodes. The conductance of moist ER materials, or ER materials with mobile ions, limits their potential use. In addition, over time, hydrous particles settle and pack into sedimentary layers.

Discoveries in the late 1980s (Ingrasia, 1988) of anhydrous electrorheologically active materials led to a significant increase in ER research, and portended the advent of a new and potentially lucrative industry. The particles in these modern materials are typically dried zeolites or amorphous alumino-silicates (Fili-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the International Mechanical Engineering Congress and Exposition, the Winter Annual Meeting of The ASME, Atlanta, GA, Nov. 17-22, 1996.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 10, 1995; final revision, Mar. 1, 1996. Associate Technical Editor: S. A. Berger.

Paper No. 96-WA/APM-5.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

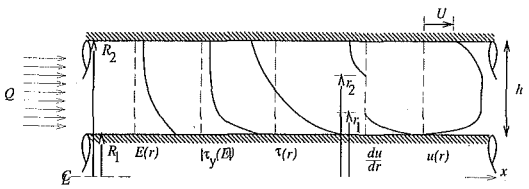


Fig. 1 Poiseuille flow of ER materials through a concentric annulus

sko, 1990). Dozens of technical and popular articles have appeared in the last five to ten years describing the potential uses of ER materials (Gavin, 1994).

The major ER effects can be qualitatively modeled by a Zener element in series with a solid-friction element and a viscous element. At high electric fields and low (pre-yield) levels of strain, ER materials are essentially viscoelastic plastic materials (Gamota and Filisko, 1991a, b).

In viscous-fluid vibration-damping devices, the flow of the fluid in the device is prescribed by the device motion. If the device has a dash-pot configuration, the fluid shear rates in the device can become quite high (on the order of a thousand per second). If this is the case, the pre-yield material behavior is of secondary importance. The observation of yield stresses motivated the use of the *Bingham* constitutive model for the fluid, in which an applied shear stress, τ, is resisted by a yielding component, τ_y, and a viscous component, $\eta\dot{\gamma}$

$$\tau(E, \dot{\gamma}) = \tau_y(E)\ \text{sgn}\ \dot{\gamma} + \eta\dot{\gamma}. \tag{1}$$

Assumptions implicit in the Bingham approximation are

- the material has no pre-yield deformation, and
- the flow of the material is steady and fully developed.

The Newtonian viscosity, η, is roughly independent of the applied electric field. The Bingham model does not account for all the interactions and mechanisms in actual ER materials. For example, it does not describe the complex behavior of ER materials as the shear rate passes through zero. Electrophoresis and nonuniform electric fields can lead to nonhomogeneous material behavior, which is not modeled by Eq. (1). In dash-pot style ER devices, where shear rates can be high, the Bingham equation models the most important mechanisms. In this paper the Bingham equation is used to derive simple approximations describing the flow of ER materials through rectangular and annular ducts. While these approximations do not capture details of the behavior of actual ER devices, they are sufficiently accurate for design calculations.

2 Analysis of ER Dampers

Most ER damping devices dissipate energy by forcing fluid through narrow ($0.5\ \text{mm} < h < 3\ \text{mm}$) laminar ducts, energized with high electric fields ($1\ \text{kV/mm} < E < 6\ \text{kV/mm}$). Using the Bingham model, an analysis of the steady Poiseuille flow of ER materials through rectangular ducts leads to simple approximations which are useful in designing ER dampers.

The flow field in the ducts of damping devices can be categorized as shear-mode, flow-mode, or a combination of shear and flow (Lou et al., 1994). Velocity fields in dash-pot ducts is caused by the rectilinear motion of the dash-pot plunger. In such dampers, the fluid velocity field always has a nonzero flow component, and sometimes has a shear component as well.

2.1 Poiseuille Flow of Electrorheological Fluids. For incompressible, generalized Newtonian fluids, in quasi-steady one-dimensional flow, the pressure gradient, $\partial p/\partial x$, along the flow, u, is resisted by fluid shear stresses, τ, which are related to the dissipative mechanisms (viscous or otherwise). In ER dampers, energy is dissipated through viscous and yielding mechanisms in the ER fluid. The yield stress is related to the controllable electric field. Figure 1 describes the flow and electric fields of an ER material flowing steadily through an annular duct. The outer electrode is moving with a velocity U with respect to the inner electrode. Assuming no net electrical charge between the electrodes, no discontinuity in dielectric properties between the electrodes, and that the potential, ε varies with r only, the electric field is found from, $E(r) = -\nabla\varepsilon$, where the voltage potential ε satisfies the Laplacian $\nabla^2\varepsilon = 0$ with the appropriate boundary conditions.

$$E(r) = \frac{\varepsilon(R_2) - \varepsilon(R_1)}{\log\,(R_2/R_1)}\frac{1}{r}, \tag{2}$$

where the scalar E is the magnitude of the field in the radial direction. The yield stress is assumed to follow a power law with the field

$$\tau_y(E, r) = \alpha|E(r)|^n, \tag{3}$$

where α and n are empirical material constants. The shear stress, $\tau(r)$, satisfies Navier-Stokes,

$$\tau(r) = -\frac{1}{2}p'r + \frac{D}{r}, \tag{4}$$

where D is a constant to be determined, and $p' = -(\partial p/\partial x)$. Under steady flow conditions, there are two coordinates, r_1 and r_2, at which $|\tau(r)| = |\tau_y(E, r)|$. Assuming ideal Bingham plastic behavior, the shear rate is zero in the region $r_1 < r < r_2$. Note that

$$r_1^2 = r_2^2 = \frac{1}{2}\frac{R_2^2 - R_1^2}{\log\,(R_2/R_1)} \tag{5}$$

and that

$$D = \frac{1}{\log\,(R_2/R_1)}\left[\eta U + \frac{p'}{4}(R_2^2 - R_1^2)\right] \tag{6}$$

if $\varepsilon_1 = \varepsilon_2$. The plug flow coordinates r_1 and r_2, and the unknown constant, D, are found numerically (Atkin, 1991) by solving $\tau(r_1) = \tau_y(r_1)$ and $\tau(r_2) = \tau_y(r_2)$ for D and $u(r_1) = u(r_2)$, with the appropriate continuity and no-slip conditions. The resulting equations are

$$\frac{r_1}{r_2} = \frac{\alpha(C/r_2)^n + p'r_2/2}{\alpha(C/r_1)^n + p'r_1/2}, \tag{7}$$

and

$$\frac{1}{\log\left(\dfrac{r_1R_2}{r_2R_1}\right)}\left[\frac{p'}{4}(R_2^2 - R_1^2 + r_1^2 - r_2^2) + \eta U + \frac{\alpha C^n}{n-1}(R_2^{1-n} + R_1^{1-n} - r_2^{1-n} - r_1^{1-n})\right] = \frac{p'}{2}\frac{r_1^{n+1} + r_2^{n+1}}{r_1^{n-1} + r_2^{n-1}} = D, \tag{8}$$

where $C = |\varepsilon_2 - \varepsilon_1|/\log\,(R_2/R_1)$. The volumetric flow rate, Q, is given by

$$Q = \frac{\pi}{8\eta}\left[p'(R_2^4 - R_1^4 + r_1^4 - r_2^4) - 4D(R_2^2 - R_1^2 + r_1^2 - r_2^2) - 8\frac{\alpha C^n}{3-n}(R_2^{3-n} + R_1^{3-n} - r_2^{3-n} - r_1^{3-n})\right]. \tag{9}$$

For flow in all dash-pot style devices, the ratio of the volumetric flow rate to the wall velocity, Q/U, is a positive constant related to the geometry of the device (Gavin, 1994).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

2.2 Nondimensional Solutions for Rectangular Ducts. The analysis in the previous section can be completed analytically if the duct section is rectangular. Phillips (1969) cast the Bingham flow equations for a rectangular duct into dimensionless forms. A set of dimensionless variables which provide some insight into the behavior of ER dash-pots is

$$P = \frac{p'}{p'_N} = \frac{bh^3 p'}{12Q\eta} \tag{10}$$

$$T = \frac{\tau_y}{hp'_N} = \frac{bh^2\tau_y}{12Q\eta} \tag{11}$$

$$V = \frac{Q_s}{Q} = \frac{bhU}{2Q} \tag{12}$$

$$A = P - 2T \tag{13}$$

where p'_N is the pressure gradient of a Newtonian fluid ($\tau_y = 0$) responding to the volumetric flow rate, Q, of the Bingham flow, and Q_s is the volume flow rate for pure shear flow. The rectangular duct has dimensions $b \times h (h \ll b)$. The dimensionless pressure gradient, P, is the ratio of the pressure gradient of a flowing ER material with a yield stress, τ_y, to the pressure gradient of a flowing ER material with no yield stress. The dimensionless wall velocity, V, is the ratio of a pure-shear flow rate to the total (shear and pressure driven) flow rate. And A is the dimensionless excess pressure gradient required to overcome the yield stress in a Bingham fluid.

The dimensionless pressure gradient, P, has physical significance as the dynamic range of a flowing ER material. It gives the ratio of the maximum damper force to the minimum damper force, at the same damper velocity. Because damper forces are, essentially, proportional to the pressure gradient through the ducts, ER dampers can be designed to meet prespecified ranges of controllable forces. In some applications it may be desirable to have precise control of the damper forces over a relatively small, but predefined, range. To take full advantage of the variable yielding properties of ER materials, it is often desirable to design devices which will exhibit large changes in behavior when the electric field is applied or removed. In either case, the damper's objectives can be achieved by designing devices to operate at the desired values of P.

Using these dimensionless variables, and simplifying equations of continuity and equilibrium, the fifth-degree polynomial

$$P^3 - (1 + 3T)P^2 + 4T^3 + P^2V + \frac{P^2TV^2}{3A^2} = 0 \tag{14}$$

describes the Poiseuille flow of Bingham materials in a rectangular duct (Phillips, 1969). Equation (14) describes the flow in all dash-pot style dampers with rectangular ducts (Gavin, 1994). This flow condition results in a detached plug between the electrodes, regardless of the (steady) wall velocity.

If $U = 0$, then $V = 0$, and (14) becomes a cubic in P, with three real roots. The desired root of this cubic is (Gavin, 1994)

$$P(T) = \frac{2}{3}(1 + 3T)\left[\cos\left(\frac{1}{3}\arccos\left(1 - 54 \times \left(\frac{T}{1 + 3T}\right)^3\right)\right) + \frac{1}{2}\right]. \tag{15}$$

A uniform approximation to (15) is accurate to within ± two percent over most of the interval $0 < T < 1000$.

$$P(T) = 1 + 2.07T + \frac{T}{1 + 0.4T} \pm \text{two percent} \tag{16}$$

When $U > 0$ the polynomial is fifth degree, and no general solution exists for the roots. For dash-pot devices, constraints on the possible flow conditions ($QU > 0$) bound V to the interval (0, 0.5), and simplify the solution to this equation (Gavin, 1994). The expression

$$P(T, V) = 1 + 2.07T - V + \frac{T}{1 + 0.4T} - \frac{1.5TV^2}{1 + 0.4T^2} \pm \text{three percent} \tag{17}$$

approximates the desired root of (14) in these cases and reduces to (16) if $V = 0$. The approximation (17) is plotted as the solid line ($V = 0$) and as a dashed line ($V = 0.5$) in Fig. 2. The dynamic range increases roughly linearly with the dimensionless yield stress, and approaches 1 for Newtonian materials ($\tau_y = 0$). The relative error of the approximation (16) to the exact Eq. (15) is shown as the solid line in Fig. 4.

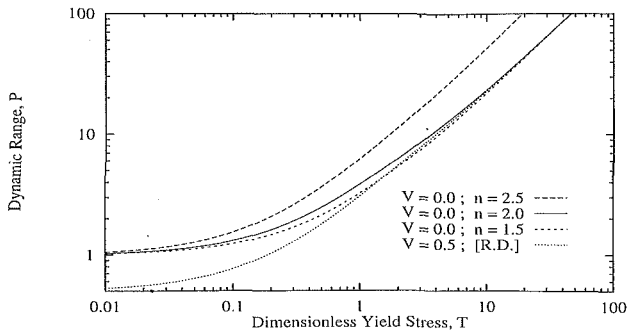


Fig. 2 Influence of the power law exponent, n, and the dimensionless yield stress, T, on the dynamic range, P, of an annular ER damper. And influence of the nondimensional wall velocity on the dynamic range of a rectangular ER damper [R.D]

These relations are useful in simulating the behavior of flowing Bingham materials, and in designing devices based on steady flow. Given material properties, τ_y and η, the duct geometry, b and h, the flow rate, Q, and a wall velocity U, (i.e., a trial device design and operating conditions) (17) determines the responding pressure gradient (and device forces) in closed form. Pressure gradient measurements in several Poiseuille flow geometries have verified that (14) is representative of actual ER flows to within five percent (Brooks, 1991; Coulter and Duclos, 1989; Duclos, 1988; Gavin, 1994; Weiss, 1993). Hence (17) can be used to design ER devices in which the flow is steady.

The gap size, h, is an important design variable that has a great influence on the dynamic range, P. Enlarging the gap in a dash-pot increases the dynamic range by reducing both the volumetric flow rate and the viscous stresses. Both P and T depend on the gap size, h, and on the flow rate, Q. So, given ER material properties, the dynamic range of an ER duct is a function of both h and Q. For rectangular ducts, this is illustrated in the solid lines of Fig. 3. Devices with wider gaps and lower flow rates have larger dynamic ranges. The dynamic range increases exponentially with a linearly increasing gap.

2.3 Application to Annular Ducts. The approximate solution to the nondimensional forms of the Bingham steady flow equations for rectangular ducts is very useful for simulation and design. Similar equations can not be derived analytically for annular ducts, due to the nonhomogeneous electric field, yield stress, and nonuniform shear stress gradient. Furthermore, the annular duct solution depends on material constants α and n. The flow of ER materials through annular ducts, can nevertheless be solved numerically (using the procedure outlined above). The dimensionless pressure gradient and yield stress can be evaluated, and closed-form expressions can be found to

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

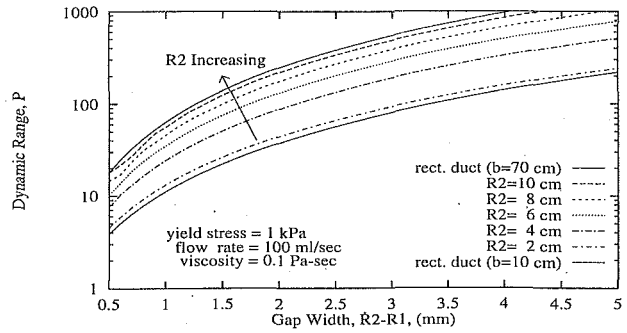


Fig. 3 Effect of gap width, $h = R_2 - R_1$, and outer radius diameter, R_2, on the dynamic range of annular ER ducts, with a comparison to rectangular ducts

approximate the numerical solution. The dimensionless pressure gradient and yield stress for annular ducts is defined as in (11) and (12), where

$$p'_N = \frac{8\eta Q}{\pi}\left[R_2^4 - R_1^4 - \frac{(R_2^2 - R_1^2)^2}{\log\frac{R_2}{R_1}}\right]^{-1}, \qquad (18)$$

and $h = R_2 - R_1$.

The effect of the power law exponent, n, on the dynamic range, P, of an ER device is illustrated in Fig. 2. The dynamic range is enhanced by a factor of two for $n = 2.5$, but is not significantly degraded for $n = 1.5$. The dynamic range, P, depends on the device geometry as illustrated in Fig. 3, assuming that the same ultimate τ_y is achieved for all the gap spacings. If $n = 2$, the dynamic range of an annular duct is similar to that of a rectangular duct with a duct width, b, equal to $\pi(R_1 + R_2)$.

To find an expression similar to (16) for annular ducts, the ER flow equations for an annulus were solved numerically for a variety of values of (R_1/R_2), n, α. For $n = 2$, the numerical solution for the annular ER flow equations matched the analytical solution for the rectangular duct (15) exactly, regardless of (R_1/R_2) and α. Furthermore, no correction need be made for (R_1/R_2), regardless of R_1, R_2, α, or n. Therefore, an approximation was chosen which reduces to (16) when $n = 2$. The following equation approximates (to within ±5 percent) the results of numerical solutions for $1.5 \le n \le 2.5$.

$$P = 1 + [2.07 + 3(n-1)(n-2) + (2.5 - 1.8|n-2|)|n-2|]T + \frac{(1 - 1.5|n-2|)T}{1 + (0.4 + 4|n-2|)T} \pm \text{five percent.} \qquad (19)$$

The relative errors of this approximation to the numerical solution, along with the relative error for the rectangular duct solution are illustrated in Fig. 4. The "+" marks in Fig. 4 are

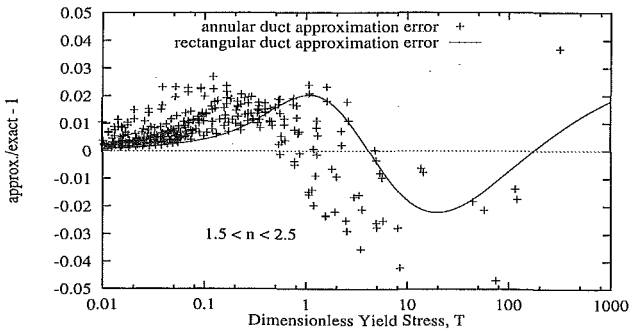


Fig. 4 Relative error of the ER flow approximation for annular (+) and rectangular (−) ducts ($1.5 < n < 2.5$)

approximation errors corresponding to a wide range of radius ratios (R_1/R_2), material properties, and flow rates. High electric fields require narrow ducts in ER devices, so the radius ratio (R_1/R_2) will be near 1 for most ducts, and the approximations for rectangular ducts will model the flow very well, regardless of the exponent, n.

2.4 The Bingham-to-Newtonian Flow Transient. The effects of time-varying material properties, sinusoidal wall motions, and flow development at the duct inlet, perturb fully developed flows (Lou et al., 1994). For control applications, an important flow transient is the change from a nearly plugged Bingham flow to a Newtonian flow. The initially rectangular ER flow profile asymptotically approaches a Newtonian profile when the electric field is instantaneously removed. The yield stress is assumed to drop to zero much more quickly than the dissipation rate of the ensuing flow transient.[1] Because the yield stress is assumed to vanish instantaneously, the development of the transient flow profile is governed by Newtonian fluid mechanics,

$$p' + \frac{\eta}{r}\frac{\partial u}{\partial r} + \eta\frac{\partial^2 u}{\partial r^2} = \rho\frac{\partial u}{\partial t}, \qquad (20)$$

and is subject to no-slip boundary conditions at the electrodes and an assumed initial flow profile,

$$u(r, 0) = U_0\left[1 - \left(\frac{2r - a - 1}{a - 1}\right)^p\right], \qquad (21)$$

where r is normalized by R_2, $(a = (R_1/R_2) \le r \le 1)$, and the (even) exponent p is much greater than one. (A value $p = 20$ was chosen for this analysis.) The associated homogeneous partial differential equation is

$$\frac{\rho}{\eta}\frac{\partial u}{\partial t} = \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial r^2}. \qquad (22)$$

This mixed, B.V.P.–I.V.P. can be solved by assuming that the solution, $u(r, t)$, can be separated in the form $\mathcal{R}(r)\mathcal{T}(t)$, which leads to two ordinary differential equations:

$$\mathcal{T}' + \beta^2\frac{\eta}{\rho}\mathcal{T} = 0, \qquad (23)$$

and

$$\mathcal{R}'' + \frac{1}{r}\mathcal{R}' + \beta^2\mathcal{R} = 0, \qquad (24)$$

which is the Bessel equation. Analyzing these ordinary differential equations, the solution to (22) subject to the no-slip boundary conditions is

$$u_h(r, t) = \sum_{j=1}^{\infty} D_j e^{-\beta_j^2\eta t/\rho}[Y_0(\beta_j a)J_0(\beta_j r) - J_0(\beta_j)Y_0(\beta_j r)], \qquad (25)$$

where β_j satisfies the characteristic equation

$$\Delta(\beta_j) = J_0(\beta_j a)Y_0(\beta_j) - Y_0(\beta_j a)J_0(\beta_j) = 0, \qquad (26)$$

and $\mathcal{R}(r)$ is the term in the square brackets of (25). Equation (26) is solved for β_j using Newton-Raphson iterations and the following approximation for the derivative:

[1] Close range repulsive forces between particulates in the ER suspension contribute significantly to the fast disruption of structure (and the associated drop in yield stress) commonly observed in ER suspensions.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

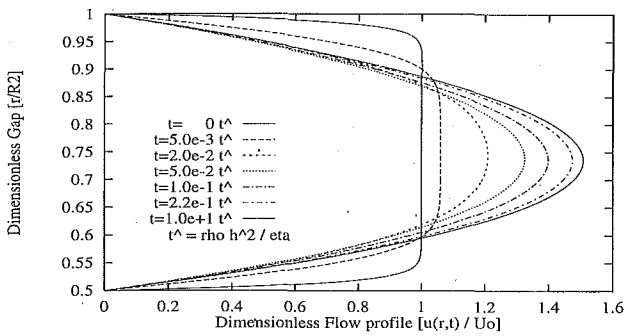


Fig. 5 Flow transient from Bingham to Newtonian behavior in an annulus ($a = (R_1/R_2) = 0.5$)

$$\Delta'(z) \approx \frac{0.8}{z}\left[\cos((1-a)z)(1-a) - \frac{1}{z}\sin((1-a)z)\right]. \quad (27)$$

The particular solution to the partial differential equation is the fully developed flow profile of a Newtonian fluid through an annulus.

$$u_p(r,\infty) = 2U_0(1-a^2) \times \frac{a^2 - r^2 - (1-a^2)[1 + \log(r)/\log(a)]}{1 - a^4 + (1-a^2)^2/\log(a)}. \quad (28)$$

The mean flow velocity is

$$U_0 = \frac{p'}{8\eta(a^2-1)}\left[\frac{1 - a^4 - (1-a^2)^2}{\log(a)}\right]. \quad (29)$$

Now the homogeneous and particular solutions are combined: $u(r, t) = u_h(r, t) + u_p(r, \infty)$. Substituting (25) into $u(r, t)$, setting $t = 0$, and rearranging, leads to

$$u(r,0) - u_p(r,\infty) = \sum_{j=1}^{\infty} D_j \mathcal{R}(r). \quad (30)$$

The coefficients D_j are then obtained using the Fourier-Bessel Theorem (Bateman, 1953).

$$\{[J_0(\beta_j a)]^2 - [J_0(\beta_j)]^2\} D_j = \tfrac{1}{2}\pi^2\beta_j^2[J_0(\beta_j a)]^2 \times \int_a^1 [u(r,0) - u_p(r,\infty)]\mathcal{R}(r) r dr, \quad (31)$$

where $u(r, 0)$ is assumed and $u_p(r, \infty)$ is known.

The Fourier-Bessel integral (31) was solved numerically using 1000 integration points. One hundred terms were retained in the series expansion (25). Then these values for D_k were substituted into $u(r, t)$ and flow profiles were calculated at various times. The flow velocity was nondimensionalized by dividing by U_0, and time was nondimensionalized by dividing by a characteristic time, $\hat{t}$, $\hat{t} = \rho(1-a)^2R_2^2/\eta$. The nondimensional flow transient is shown in Fig. 5. Values for the geometry parameter, a, of 0.5, 0.7, and 0.9 were investigated. The case $a = 0.5$ is shown, as it is the most unlike a rectangular duct. In all cases, the maximum flow velocity reaches 0.707 of its final velocity in $0.019\hat{t}$ and 0.99 of its final velocity in $0.22\hat{t}$. For typical materials, $\hat{t}$ can range from 1 ms ($h = 1$ mm) to 10 ms ($h = 3$ mm).

Like the dynamic range, P, $\hat{t}$ increases with h^2. So, $\hat{t}/\rho$ can be substituted for h^2/η in (16). An example of the interdependence of P and $\hat{t}$ is illustrated in Fig. 6 (for devices where $V \approx 0$). The tradeoff between dynamic range and response time can be controlled by the gap size h. Dampers using flowing ER materials can not simultaneously have a short response time and a high dynamic range. Increasing h to achieve a larger dynamic range is done at the expense of slower transients. For gaps less than 1 mm, the speed of the material property transition, not inertia, dominates the total fluid response time.

The response time depends on the nature of the ER effect as well (Boyle, 1991). This analysis is contingent on fast material response times. Because the longer time-scales associated with coarsening of fibrils, material response times are longer for antithixotropic changes than for thixotropic changes (high τ_y to low τ_y). ER devices described in the literature commonly have dynamic ranges of 5–15 at frequency bandwidths of 100 Hz–300 Hz (Gavin, 1994).

There are limiting factors which control the range of possible gap sizes. Coagulating particles can block small gaps even at $E = 0$. Gap sizes are limited by the voltages required to produce electric fields of 3–7 kV/mm (Bullough, 1988). As the device voltage approaches 20 kV, X-rays are produced. In practice, it is easier to apply higher electric fields through larger gaps because impurities suspended in the material do not bridge wider spaced electrodes as frequently. Hence, gaps in ER devices should range from 0.2 to 3.0 mm.

3 Design of ER Dampers

Important device characteristics specified by a particular application are: the force-velocity relationship with no electric field, the force-velocity relationship with high electric field, the time required for the device to respond to a step change in electric field, and the size of the device. Given a trial device design, its operating velocity ranges, and ER material properties, device performance characteristics can be approximated using the relations presented in this paper.

An ER damper can be designed using the principle of a dash-pot, where the piston head of the dash-pot is replaced by a rigidly interconnected array of parallel (or concentric) electrodes of alternating polarity. As the rigidly interconnected plates translate rectilinearly through the fluid, they pump fluid from one side of the ER damper, through the gaps between the plates, to the chamber on the other side of the plates. Assuming incompressible flow, the mean flow velocity, V_f, through the $N + 1$ rectangular gaps between N plates and the case is

$$V_f = \left(1 + \frac{N}{N+1}\frac{t}{h}\right)U, \quad (32)$$

where h is the gap size, t is the thickness of the plates, and U is the device velocity. Balancing shear and pressure forces on the plates in the device leads to the force, F, as a function of the pressure difference between the two chambers, $p'l$,

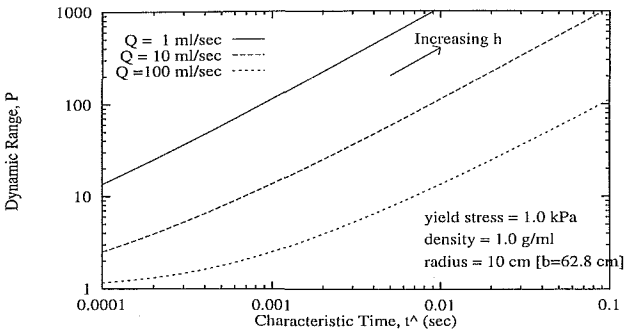


Fig. 6 Dynamic range, P, as a function of response time, $\hat{t}$, and flow rate, Q

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$F = Np'l(t + h)b, \quad (33)$$

where l and b are the length and the width of the plates, respectively. This expression assumes equal flow profiles in *all* the gaps. This assumption does not lead to large errors in most cases. This is due to the fact that $P(T, V)$ is close to $P(T, 0)$ for $T > 1$ and $V < 0.2$ (Gavin, 1994). The pressure gradient, p', can be calculated by substituting $V_f bh$ for Q in (16). Different values for p' will be obtained for high and low electric fields. In addition, p' depends on the device velocity. Since (33) depends on p' and device geometry, the device force can be calculated directly from pressure gradient measurements. Pressure measurements and (33) are used to calculate device forces in experiments on ER dampers (Gavin et al., 1995). Calculating the force in this fashion eliminates the friction caused by seals and other sliding mechanisms within the device; this force is therefore representative of the ER fluid behavior.

4 Summary

Electrorheological (ER) materials exhibit remarkable changes in material properties when subjected to strong electric fields (kV/mm). Primitive ER materials suffer from low levels of controllable stress, high current densities (kW of electrical power for kN of mechanical force (Klass and Martinek, 1967)), arcing, and irreversible sedimentation of the particulates over time. The effect of constitutive material properties such as: particle size, morphology, stoichiometry, ion concentration, permittivity, and dielectric loss need further study in order to develop better models for the ER effect and processing procedures for strong and reliable ER materials (Filisko, 1991; D.o.E., 1993). More recent anhydrous zeolite ER materials draw negligible currents (10^{-12} Amp/cm^2). Better mechanical properties accompany these improved electrical properties.

Viscoelastic behavior is the dominant characteristic of ER materials at low strains, and a yielding mechanism dominates ER behavior at shear rates greater than $\approx$500/sec. In steady, fully developed flow, the behavior of ER materials approaches that of a Bingham fluid, in which an applied shear stress, τ, is resisted by a yielding component, τ_y, and a viscous component, $\eta\dot{\gamma}$. While oscillating flows violate these assumptions, the Bingham model's simplicity makes it attractive for rough calculations of the force capacity of ER dash-pots.

Despite many material complexities, the Bingham constitutive model for ER fluids has been verified experimentally for a number of steady flow configurations (Brooks, 1991; Coulter and Duclos, 1989; Duclos, 1988; Gavin, 1994; Weiss, 1993) and seems to be sufficiently accurate for design purposes.

In this paper the Navier-Stokes equations were solved for steady flow of an electrorheological fluid through annular ducts. An approximation to nondimensionalized numerical solutions is presented. These approximations are very useful in designing ER devices. These analyses predict that the dynamic force range of dash-pot devices increases exponentially with the ratio of the gap size to plunger diameter.

The solution to the flow transient due to the changing material properties can be used to develop performance functions relating the dynamic range of an ER device and its response time to the dimensions of the device, and the properties of the ER material. Increasing the dynamic range by increasing the gap size is afforded at the expense of the device's response times. Thus the design of ER devices entails the balancing of the dynamic range with the response time.

The simplicity of device designs using ER materials (no moving parts, direct transduction from a command signal to a control force) and the fast response times (1–10 ms) makes them attractive for vibration control applications (McClamroch et al., 1994). Desirable properties of the ER devices for such applications (fast response times and large dynamic range) are evident and can be assessed using the relationships described herein.

Magnetorheological materials behave, qualitatively, very similarly to ER materials, the major difference being that magnetorheological materials are suspensions sensitive to magnetic fields. The relations presented in this paper are equally applicable to both ER and MR device design.

Acknowledgments

This material is based on work supported by the National Science Foundation under Award No. BCS-9201787 as part of the Coordinated USA Research Program on Structural Control for Safety, Performance, and Hazard Mitigation. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

The author is thankful to the referees for their insightful and helpful comments, and to the editors for coordinating the review of this paper.

References

Atkin, R. J., Shi, X., and Bullough, W. A., 1991, "Solution of the constitutive equations for the flow of an electrorheological fluid in the radial configuration," *Journal of Rheology,* Vol. 35, No. 7, pp. 1441–1461.

Bateman, H., 1953, *Higher Transcendental Functions,* Vol. II, McGraw-Hill, New York, pp. 70–72.

Block, H., and Kelly, J. P., 1988, "Review Article: Electro-rheology," *Journal of Physics D: Applied Physics,* Vol. 21, pp. 1661–1677.

Block, H., Kelly, J. P., Qin, A., and Watson, T., 1990, "Materials and Mechanisms in Electrorheology," *Langmuir,* Vol. 6, pp. 6–14.

Bonnecaze, R. T., and Brady, J. F., 1992, "Dynamic simulation of an electrorheological fluid," *Journal of Chemical Physics,* Vol. 96, No. 3, pp. 2183–2202.

Boyle, F. P., 1991, "Performance Characterization of ER Fluids: Durability," *Proceedings of the International Conference on Electrorheological Fluids,* Oct. 15-16, Carbondale, IL, World Scientific Press, Singapore, pp. 236–245.

Brooks, D., Goodwin, J., Hjelm, C., Marshal, L., and Zukoski, C., 1986, "Visco-Elastic Studies on an Electro-Rheological Fluid," *Colloids and Surfaces,* Vol. 18, pp. 293–312.

Brooks, D. A., 1991, "Design And Development Of Flow Based Electro-Rheological Devices," *Proceedings of the International Conference on Electro-rheological Fluids,* Oct. 15-16, Carbondale, IL, World Scientific Press, Singapore, pp. 367–397.

Bullough, W. A., and Peel, D. J., 1988, "Progress Towards a Hydraulic Semiconductor for Vehicle Applications," SAE Technical Paper Series 881786, pp. 1–15.

Chen, Y., Sprecher, A. F., and Conrad, H., 1991, "Electrostatic particle-particle interactions in electrorheological fluids," *Journal of Applied Physics,* Vol. 70, No. 11, pp. 6796–6802.

Conrad, Hans, and Sprecher, Arnold F., 1991, "Characteristics and Mechanisms of Electrorheological Fluids," *Journal of Statistical Physics,* Vol. 64, No. 5-6, pp. 1073–1091.

Coulter, J. P., and Duclos, T. G., 1989, "Applications of Electrorheological Materials in Vibration Control," *Proceedings of the 2nd International Conference on ER Fluids,* Aug. 7-9, Raleigh, NC., Technomic Publishing Co., Lancaster, pp. 300–325.

Davis, L. C., 1993, "The metal-particle/insulating oil system: An ideal electrorheological fluid," *Journal of Applied Physics,* Vol. 73, No. 2, pp. 680–683.

Department of Energy, 1993, "Electrorheological (ER) Fluids: A Research Needs Assessment Final Report," DOE/ER/30172 May 1993 Office of Scientific and Technical Information, Oak Ridge, TN.

Duclos, T. G., 1988, "Design of Devices Using Electrorheological Fluids," SAE Technical Paper Series 881134, pp. 1–5.

Filisko, F. E., and Radzilowski, L. H., 1990, "An intrinsic mechanism for the activity of alumino-silicate based electrorheological materials," *Journal of Rheology,* Vol. 34, No. 4, pp. 539–552.

Filisko, F. E., 1991, "Materials in Electrorheology," Technical report, Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI.

Filisko, F. E., and Wang, H-R, 1994, "The Dynamic Character of ER Materials," *Polymer Preprints,* Vol. 35, No. 2, pp. 329–330.

Gamota, D. R., and Filisko, F. E., 1991, "Dynamic mechanical studies of electrorheological materials: Moderate frequencies," *Journal of Rheology,* Vol. 35, No. 3, pp. 399–425.

Gamota, D. R., and Filisko, F. E., 1991, "High frequency dynamic mechanical study of an aluminosilicate electrorheological material," *Journal of Rheology,* Vol. 35, No. 7, pp. 1411–1425.

Gavin, H. P., 1994, "Electrorheological Dampers for Structural Vibration Suppression," Ph.D. Dissertation, The University of Michigan, Ann Arbor, MI.

Gavin, H. P., Hanson, R. D., and McClamroch, N. H., 1995, "Electrorheological Dampers, Part II: Testing and Modeling," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 676–682.

Halsey, T. C., and Martin, J. E., 1993, "Electrorheological Fluids," *Scientific American,* Oct., pp. 58–64.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Ingrasia, P., 1988, "Michigan Scientist To Receive Patent for 'Smart Fluids'," *The Wall Street Journal,* May 17, p. 58.

Jordan, T. C., and Shaw, M. T., 1989, "Electrorheology," *IEEE Transactions on Electrical Insulation,* Vol. 24, No. 5, pp. 849–878.

Klass, D. L., and Martinek, T. W., 1967, "Electroviscous Fluids. II. Electrical Properties," *Journal of Applied Physics,* Vol. 38, No. 1, pp. 75–80.

Klingenberg, D. J., van Swol, F., and Zukoski IV, C. F., 1991, "The small shear rate response of electrorheological suspensions. II. Extension beyond the point-dipole limit," *Journal of Chemical Physics,* Vol. 94, No. 9, pp. 6170–6178.

Klingenberg, D. J., Zukoski IV, C. F., and Hill, J. C., "Kinetics of structure formation in electrorheological suspensions," *Journal of Applied Physics,* Vol. 73, No. 9, pp. 4644–4648.

Lou, Z., Ervin, R. D., and Filisko, F. E., 1994, "A Preliminary Parametric Study of Electrorheological Dampers," ASME *Journal of Fluids Engineering,* Vol. 116, pp. 570–576.

Marshall, L., Zukoski IV, C. F., and Goodwin, J. W., 1989, "Effects of Electric Fields on the Rheology of Non-aqueous Concentrated Suspensions," *Journal of the Chemical Society, Faraday Transactions 1,* Vol. 85, No. 9, pp. 2785–2795.

McClamroch, N. H., Ortiz, D. S., Gavin, H. P., and Hanson, R. D., 1994, "Electrorheological Dampers and Semi-Active Structural Control," *Proceedings of the 33rd IEEE Conference on Decision and Control,* Orlando, FL.

Phillips, R. W., 1969, "Engineering Applications of Fluids with a Variable Yield Stress," Ph.D. Dissertation, Department of Mechanical Engineering, University of California, Berkeley, CA.

Schubring, A. W., and Filisko, F. E., 1994, "Study of the Variation of the Substitutional Atoms in Amorphous Alumino-Silicates as an Electrorheological Material," *Polymer Preprints,* Vol. 35, No. 2, pp. 369–370.

Weiss, K. D., Carlson, J. D., and Coulter, J. P., 1993, "Material Aspects of Electrorheological Systems," *Journal Intelligent Material Systems and Structures,* Vol. 4, pp. 13–34.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

H. P. Gavin
Assistant Professor,
Department of Civil and
Environmental Engineering,
Duke University,
Durham, NC 27708-0287
Assoc. Mem. ASME

R. D. Hanson
FEMA Disaster Field Office,
245 S. Los Robles, Room 630,
Pasadena, CA 91102-6020

F. E. Filisko
Professor,
Department of Materials Science
and Engineering,
The University of Michigan,
Ann Arbor, MI 48109

Electrorheological Dampers, Part II: Testing and Modeling

Electrorheological (ER) materials develop yield stresses on the order of 5–10 kPa in the presence of strong electric fields. Viscoelastic and yielding material properties can be modulated within milli-seconds. An analysis of flowing ER materials in the limiting case of fully developed steady flow results in simple approximations for use in design. Small-scale experiments show that these design equations can be applied to designing devices in which the flow is unsteady. More exact models of ER device behavior can be determined using curve-fitting techniques in multiple dimensions. A previously known curve-fitting technique is extended to deal with variable electric fields. Experiments are described which illustrate the potential for ER devices in large-scale damping applications and the accuracy of the modeling technique.

1 Introduction

Electrorheological (ER) suspension are unique micron-sized particles in hydrophobic, insulating, dielectric oils (Block, 1988; D.o.E., 1993; Jordan and Shaw, 1989). ER materials develop yield stresses in the presence of strong electric fields. Upon removal of the field, material property changes are fully reversible. At high electric fields and low levels of strain, ER materials are essentially viscoelastic plastic materials (Gamota and Filisko, 1991a, b). In steady, fully developed flow, the behavior of ER materials approaches that of a Bingham fluid, in which the applied shear stress, τ, is resisted by a yielding component, τ_y, and a viscous component, $\eta\dot{\gamma}$.

$$\tau(E, \dot{\gamma}) = \tau_y(E) \operatorname{sgn} \dot{\gamma} + \eta\dot{\gamma}. \quad (1)$$

Assumptions implicit in the Bingham approximation are: (i) the material has no pre-yield deformation, and, (ii) the flow of the material is steady and fully developed. The Newtonian viscosity, η, is roughly independent of the applied electric field.

Using Eq. (1) as the constitutive relation, an analysis of the flow of ER materials through narrow laminar rectangular ducts can be carried out by writing equations of equilibrium and continuity. An approximation for the pressure gradient, p', along a narrow rectangular duct of dimension $b \times h (h \ll b)$, as a function of the flow rate, Q, wall velocity, U, and device geometry has been determined (Gavin et al., 1995),

$$P(T, V) = 1 + 2.07T - V + \frac{T}{1 + 0.4T} - \frac{1.5TV^2}{1 + 0.4T^2} \pm \text{three percent}, \quad (2)$$

where $P = p'/p'_N$ is a dimensionless pressure gradient, $T = \tau_y/(hp'_N)$ is a dimensionless yield stress, $V = Ubh/(2Q)$ is a dimensionless wall velocity which depends only on the device geometry in dash-pot style dampers, and p'_N is the pressure gradient corresponding to $\tau_y = 0$ (zero electric field). These dimensionless forms were first studied by Phillips (1969). Equation (2) approximates the root of a fifth-degree polynomial in P. The variable V is bounded by the interval (0, 0.5) in dash-pot devices (Gavin, 1994). The variable P has a physical interpretation as the dynamic range of a flowing ER material (Gavin et al., 1995). Equation (2) was derived for rectangular ducts. It is a good approximation for annular ducts as well, provided that the yield stress is proportional to the electric field squared, and V is small, or the ratio of the duct radii is close to 1. Equation (2) is therefore useful for designing a variety of ER devices in which the flow is steady. When combined with equilibrium and mass balance equations for a device, Eq. (2) relates the variable yield stress and the device velocity to the device force.

While Eq. (2) is useful for design purposes, more accurate models for ER device behavior can be determined using test data from actual devices. A method for fitting a model to experimental force response histories as a function of measured device motions and electric fields, results in accurate device-level models for use in simulation and control contexts.

Section 2 contains an analysis of experimental results which illustrate the accuracy of Eq. (2) for ER device design. Small-scale experiments on a simplified ER damper were designed for this purpose. To address the use of ER materials in large-scale vibration control applications, a larger device was built and tested. The behavior of this device was more complex, as described in Section 3, and motivated the use of a curve-fit model. Section 4 includes a discussion on fitting data by orthogonal polynomials; and a method using Chebyshev polynomials is applied to the data described in Section 3.

2 Small-Scale Tests for Design Verification

The purpose of this series of experiments was to evaluate the applicability of the analytical model derived for steady flow, Eq. (2), to cases of sinusoidal, slowly varying flow.

2.1 Dimensions, Materials, and Test Configuration. The test device consisted of a single prismatic plunger (t = 12.7 mm, l = 50.4 mm, b = 150 mm) which was actuated vertically through ER fluid in an open box. The plunger was made of 6061-T6 aluminum and had a clear anodized surface,

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the International Mechanical Engineering Congress and Exposition, the Winter Annual Meeting of The ASME, Atlanta, GA, Nov. 17-22, 1996.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 10, 1995; final revision, Mar. 1, 1996. Associate Technical Editor: S. A. Berger.

Paper No. 96-WA/APM-6.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

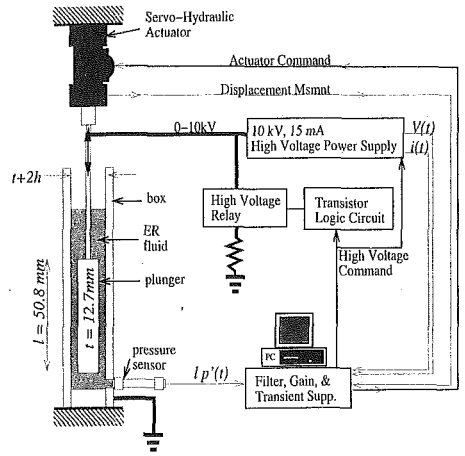


Fig. 1 Configuration of the small test device, hydraulic actuator, high voltage supply, data acquisition, and control

which sealed the metal and prevented fine aluminum particles from entering the ER material. Corners of the plunger were rounded to reduce local charge concentrations, and to provide a more streamlined duct entrance. A constant gap was maintained by Plexiglas guides. The clearance between the plunger and the box, h, could be adjusted between 1 mm and 3 mm.[1] As the plunger moved through the ER fluid, fluid was displaced and forced to flow through the gap.

Figure 1 illustrates the test configuration. Custom signal conditioning provided low-pass filtering, gain, and most importantly, high-voltage transient suppression via high-voltage metal-oxide varistors.[2] Preprogrammed displacement command signals were smoothed by a custom filter prior to controlling a servo-hydraulic actuator. The device, actuator, and PC were operated during high-voltage arcing with no discernible effect on the actuator performance or the collected data. As the plunger was cycled through the ER fluid, measurements of ER fluid pressure response, actuated plunger displacement, applied high voltage, and electrical current were collected. Tests were conducted with constant and fast-slewing electric fields. The ER test device never drew more than 5 Watts, and could have been operated from a battery.

The ER behavior of the material used in this series of tests (100 micron aluminosilicate particles in light mineral oil) was obtained using a Couette rheometer. The value of the yield stress depends strongly on the value of the shear rate at which yielding is defined (Gavin, 1994). Yield stresses decrease with increasing shear rates. This is likely due to the more rapid destruction of fibrated structure of particulates at higher shear rates. In this study, the yield stress was defined as an average of the electric field dependent shear stresses between shear rates of 0.0048/sec and 0.48/sec. With this definition, the yield stress for the ER material used in this study is related to the electric field by the power law $\tau_y(E) = 73.9E^{1.27}$, where τ_y is in Pa and E is in kV/mm. The viscosity η is roughly independent of the field, E.

2.2 Constant Voltage Tests. In the constant voltage tests, ten sinusoidal cycles at 6, 3, and 1.2-sec periods were actuated with linearly increasing amplitudes (in time). Tests were conducted at constant electric field strengths between 0 and 3.34 kV/mm. Figures 2(a) and 2(b) illustrate the repeatability and range of the hysteresis loops obtained from the constant E tests. Each of the figures illustrates four sets of hysteresis loops: small and large amplitudes, and zero and high voltage. The more rounded hysteresis loops at higher frequencies, and smaller gaps in Fig. 2(a) are due to an increase in viscous stresses. As the viscous stresses increase (smaller gap or higher velocity) the ratio of maximum to minimum forces decreases. In the long period motions, the deformation rates are much lower and the pre-yield behavior becomes important. A static yield stress, in excess of the dynamic yield stress, can be inferred from Fig. 2(b). If a maximum range of controllable energy dissipation is the design objective, then the gap size should be set such that the difference between the static and dynamic yield stress is just recovered by viscous stresses. So doing, the device will behave like an ideal variable friction damper, and will have velocity-independent, voltage-dependent, damping.

Using the measurements described above, coordinates (T, P) can be calculated from the test data, and the power law material equation.[3] These data points are plotted along with the approximate solution (2) in Figs. 3(a) and 3(b). Error bars are given for a 90 percent (Student-t) confidence interval for the mean. Note that in the tests with $h = 2.38$ mm, the dynamic range, P, is an order of magnitude greater than the dynamic range for the tests with $h = 0.96$ mm (for the same applied test motions). From these figures it is clear that the Bingham model provides an upper limit to the forces developed in ER devices. At small deformations, mechanisms not modeled by (1) dominate the behavior. In this region, the viscoelastic and pre-yield properties of the ER material are important. The presence of nonhomogeneous material properties could be another source of scatter in this data.

Measured pressure response time histories and those predicted from the Bingham model (Fig. 4) provide a different comparison of test results and Eq. (2). Note that the measured pressure response lags the theoretical response of an ideal inertia-less Bingham fluid. This transient is due to the nonsteady prescribed wall motion and flow.

[1] Gaps much larger than 3 mm require X-ray level voltages (20 kV) for fields capable of producing a strong ER effect.

[2] MOV's are two terminal passive devices which behave much like bidirectional Zener diodes.

[3] A value for η equal to 0.35 Pa was chosen based on the test results at $E = 0$.

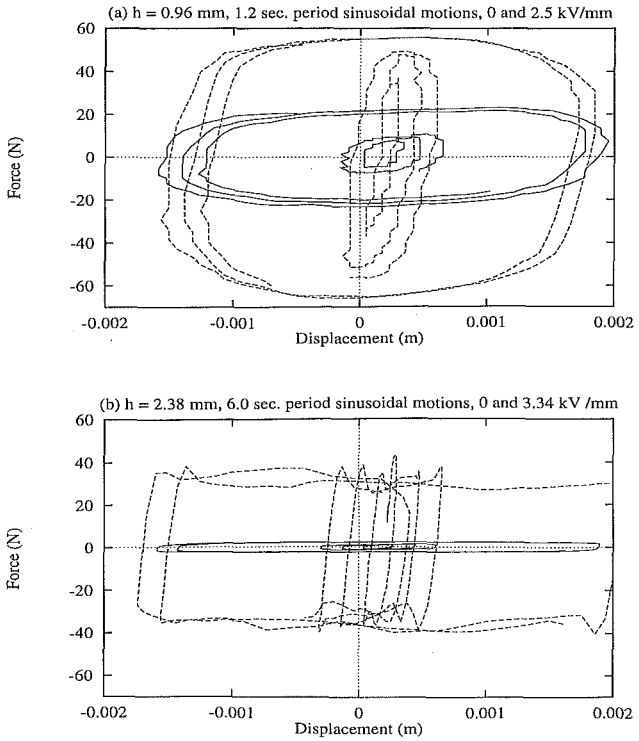


Fig. 2 Experimental ER hysteresis loops at constant voltages: (*a*) E = 0.0 and 2.50 kV/mm; 1.2 sec. sinusoidal period; and h = 0.96 mm. (*b*) E = 0.0 and 3.34 kV/mm; 6.0 sec. sinusoidal period; and h = 2.38 mm.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

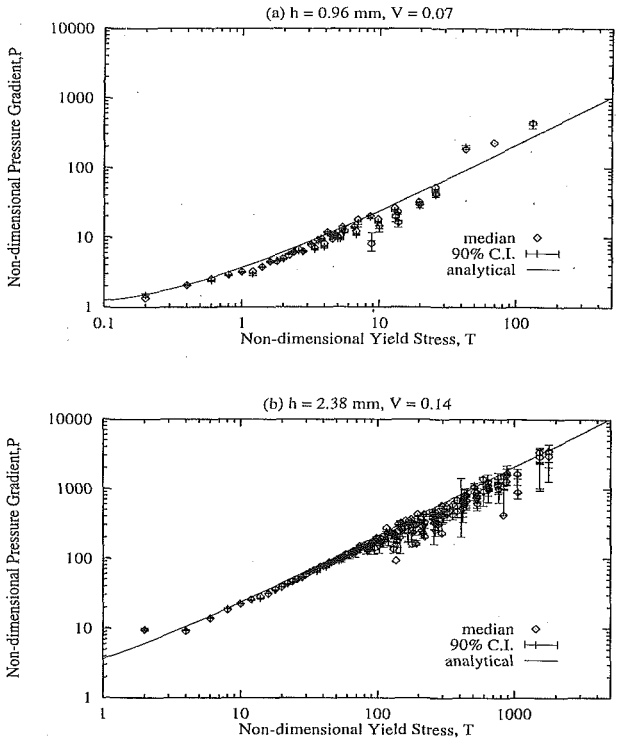


Fig. 3 Quantitative evaluation of the Bingham flow equations for ER materials: (*a*) h = 0.96 mm (*b*) h = 2.38 mm

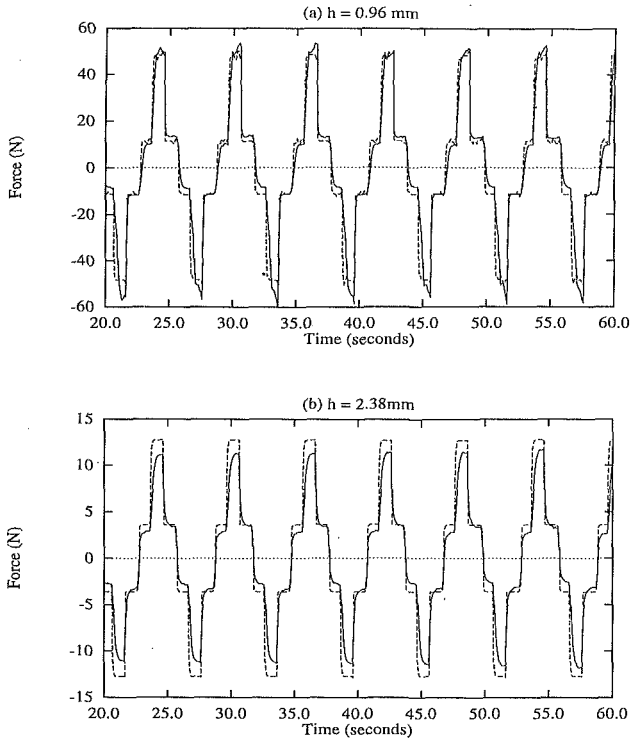


Fig. 5 Transient pressure responses for an actual ER fluid: —, and an ideal inertialess Bingham fluid: – – – –. (*a*) h = 0.96 mm, E = 2.21 kV/mm (*b*) h = 2.38 mm, E = 1.57 kV/mm.

2.3 Variable Voltage Tests. Varying electric fields also cause transients in the pressure gradient. To test the response of p' to step-like changes in E, periodic motions of constant velocity (triangle-wave displacements) were actuated in conjunction with the high voltage on the ER test device. These signals were coordinated so that pressure transients at midstroke would be due to changes in material properties alone. The RC time constant of the device depends on the (nonohmic) conductivity of the ER fluid and can range from 1 to 3 seconds.[4] Electric fields can nevertheless be discharged in micro-seconds by closing a high-voltage relay whenever the high-voltage command signal drops to zero. This control decision was implemented in an analog circuit. In Fig. 5, two sets of pressure response data (h = 0.96 mm and h = 2.38 mm) are compared to the model of (2) (which assumes that the fluid can respond arbitrarily fast and has no inertia). The salient feature of these figures is the rapid drop in force when the electric field is removed. The electrical discharge of the ER device is practically instantaneous, as is the disruption of the structured, fibrated, particulate phase; and the response time of the force is influenced primarily by fluid inertia. Force rise-times are slower, because device charging times are longer than device discharging times, and micro-structure formation is slower than micro-structure disruption. Equation (2) compares favorably to the test results for the h = 0.96 mm case, indicating that inertia effects are not important for gap sizes less than 1 mm. For the h = 2.38 mm case, Fig. 5 shows that the measured pressure response is somewhat slower than the simulated response of a fluid with zero inertia. The pressure gradient responds more slowly to increasing electric fields than to decreasing electric fields in both cases. This is primarily due to the fact that the ER material properties respond more quickly to a decreasing electric field (Gavin, 1994).

The effect of fluid inertia on the pressure response of the fluid undergoing material property changes agrees qualitatively with theoretical relations (Gavin et al., 1995). ER fluids flowing in gaps larger than 1 mm respond more slowly to changes in the yield stress, due to inertial effects. Hence, the dynamic range and response times both increase with the gap size, h.

The behavior of even simple ER devices can be quite complex. As device geometries increase in size and complexity, it is reasonable to assume that their behavior will as well. The next section describes large scale ER device characterization experiments.[5]

[4] Better ER materials conduct lower currents, and therefore have longer RC time constants.

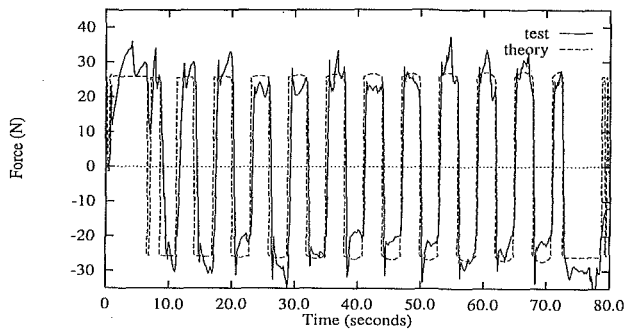


Fig. 4 Experimental ER pressure response and the pressure response of a corresponding Bingham material. (E = 2.92 kV/mm) and increasing sinusoidal amplitudes (6.0 sec. sinusoidal period, h = 2.38 mm).

3 Large-Scale Tests for Structural Applications

The purpose of the large device tests was to gain experience in the use of ER dampers for large-scale structural vibration control applications, and to generate a database for evaluating ER device modeling methods.

[5] In the course of the research, results from the large-scale tests motivated the small-scale experiments.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

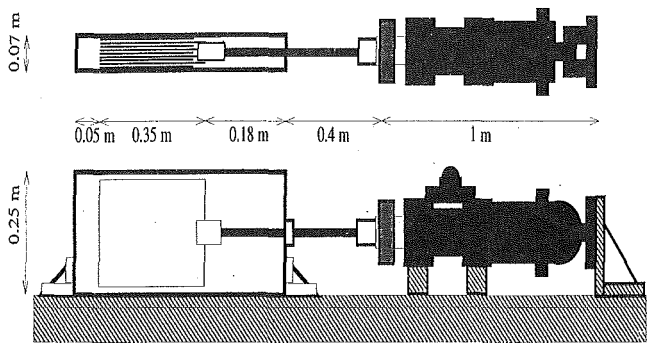


Fig. 6 Configuration of the large test device and hydraulic actuator

3.1 Dimensions, Materials, and Test Configuration. The device was comprised of an array of nine, rigidly interconnected, rectangular plates. The outer and alternate plates were connected to a high voltage potential. A $h = 0.78$ mm gap was maintained by oversized Nomex washers around the screws connecting the plates. The plates were fabricated from $t = 4.76$ mm thick 6061-T6 aluminum and measured 254 mm by 356 mm. The edges of the plates were rounded to reduce electric charge concentrations, cavitation, and the possibility of arcing. In an attempt to seal the aluminum, all surfaces exposed to ER material were anodized. Figure 6 illustrates the device and its actuator connection.

The plates moved rectilinearly within a snug-fitting case. The edges of the plates rested on, and slid along, the bottom surface of the box. To prevent short circuits of the high voltage and to reduce friction, the plate-bearing faces of the box were laminated with high dielectric strength polyester.

As the fixed plates translate rectilinearly through the fluid, they pump fluid from one side of the ER damper, through the gaps between the plates, to the chamber on the other side of the plates.

A 1-in. diameter hardened steel tubular shaft, connected to the plates, passed through the case. An O-ring energized PTFE wiper seal performed satisfactorily in sealing the device with minimum friction. The device was designed such that the plates could move with a stroke of 20 cm. A 22 gauge wire passing through the steel shaft connected the high-voltage plates to the power supply. The ground potential plates were physically and electrically connected to the shaft, which provided the (low impedance) ground connection. Plates adjacent to the case were at high voltage, thereby creating an electric field in all of the gaps. Balancing shear and pressure forces on the plates in the device leads to the force, F, as a function of the pressure difference between the two chambers, $p'l$,

$$F = Np'l(t + h)b, \tag{3}$$

where l and b are the length and the width of the plates, respectively, and N is the number of plates.

The dimensionless wall velocity, V, depends only upon the geometry of the device and was 0.08. So the pressure drop along the outer gaps is almost the same as the pressure drop along the inner gaps, and (3) is sufficiently accurate, especially in light of the unmodeled complexities of ER material behavior (Gavin, 1994).

3.2 Experimental Data. Sinusoidal motions were actuated to obtain data for characterization studies. The periods were long, reflecting the large-scale nature of the device and the envisioned application. These periods ranged from two to ten seconds. The period was constant during each test.

Initial evaluation tests were carried out using mineral oil- and kerosene-based ER fluids. The mineral oil based fluid produced force levels six times as great, but a dynamic range only half as great, as those using the kerosene-based fluid.

The ER behavior of the material used in this series of tests (1 micron alumino-silicate particles in kerosene) were, again, obtained from Couette rheometer tests. In the large-scale device tests higher shear rates than in the small scale tests were anticipated. So the yield stress was defined as the electric field-dependent shear stress at a rate of 1.42/sec. The yield stress was related to the electric field by the power law: $\tau_y(E) = 15.5E^{1.86}$, where τ_y is in Pa and E is in kV/mm. Rheometer studies of the ER fluid gave a value of 0.5 Pa-sec for η.

To obtain device data over a large portion of the phase plane, the applied sinusoidal displacements increased linearly with time. Measurements of fluid pressures across the device, actuator displacements, actuator forces, voltage levels, and electric currents were collected.

Equivalent damping ratios, c_{eq}, were calculated from the measured hysteresis loops, the measured sinusoidal period, T, and the displacement amplitude, A:

$$c_{eq} = \frac{W_D T}{2\pi^2 A^2}, \tag{4}$$

where W_D is the energy dissipated per cycle. The dependence of c_{eq} on A and E is illustrated in Fig. 7. Decreasing equivalent damping with increasing amplitude is an artifact of the Coulomb-type damping of ER devices. This reduction in c_{eq} becomes more pronounced as the electric field increases.

Figure 8 illustrates two sets of force-displacement hysteresis loops. Figure 8(a) shows simulated-force loops obtained from measured device motions, known electric fields, known ER material properties, the device geometry, and Eqs. (2) and (3). Figure 8(b) presents the measured force hysteresis loops. These loops are simply the measured forces versus the measured displacements. The Bingham equation is adequate for a rough (design-level) description of the force levels in the ER device. It cannot, however, capture the details of hysteretic behavior under cyclic loading conditions.

4 ER Device Modeling

An accurate model for the behavior of ER dampers is necessary for simulation studies of the damper behavior as well as for vibration control simulations. Some control methods require known input-output relations for the actuators providing the control forces as well. In this case, an accurate model for the device can be derived from experiments on the device itself. Test data was used to fit the device force as a unique function of displacement, velocity, and electric field.

A scalar function of many variables can be approximated, in a least-squares sense, by a series expansion of orthogonal polynomials. The coefficients of the expansion can be evaluated explicitly, provided that the orthogonality properties of the polynomials are conserved in numerical calculation of the coefficients. Of the set polynomials mutually orthogonal in a closed

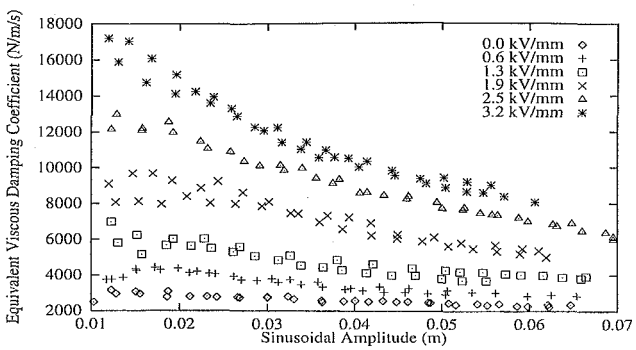


Fig. 7 Effect of amplitude of motion and electric field on the equivalent damping ratio, c_{eq}. The period of the sinusoidal motion in this test was 4.3 seconds.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

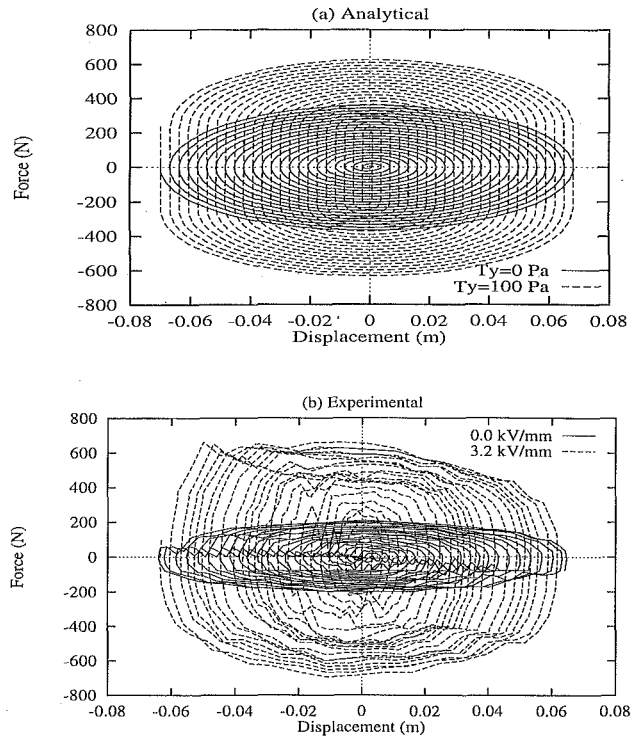


Fig. 8 (*a*) Analytical, and (*b*) experimental hysteresis loops for an electrorheological damper at low and high electric fields.

domain, Chebyshev and Legendre polynomials were considered for curve-fitting applications in this study.

4.1 Orthogonal Polynomial Approximation. Chebyshev polynomials,

$$T_n(x) = \cos\,(n \arccos\,(x)), \tag{5}$$

can be evaluated recursively (Beyer, 1984). The discrete orthogonality condition for Chebyshev polynomials is defined by

$$\sum_{k=1}^{M} T_i(x_k)T_j(x_k) = \begin{cases} 0 & i \neq j \\ M/2 & i = j \neq 0 \\ M & i = j = 0 \end{cases} \tag{6}$$

where

$$x_k = \cos\left(\frac{\pi}{M}\left(k - \frac{1}{2}\right)\right) \quad k = 1, 2, \ldots, M \tag{7}$$

are the roots of $T_M(x)$ and lie in the interval $[-1, 1](M \geq i, j)$. To take advantage of the orthogonality properties of the Chebyshev polynomials, in least-squares approximation of data, the data needs to be interpolated to the zeros of some higher order Chebyshev polynomial, $T_M(x)$.

Legendre polynomials, $P_n(x)$, can be evaluated recursively as well (Beyer, 1994). The discrete orthogonality condition for Legendre polynomials is defined by

$$\sum_{k=1}^{M} P_i(x_k)P_j(x_k) = \begin{cases} 0 & i \neq j \\ M/(2i+1) & i = j \end{cases} \tag{8}$$

where the x_k are uniformly distributed in the interval $[-1, 1]$.

The accuracy of any curve-fitting method which capitalizes upon the orthogonality of the basis functions depends implicitly upon the numerical preservation of the orthogonality conditions. The strength of Chebyshev basis curve-fits is that the orthogonality condition is accurately preserved, even if $M \sim i, j$. The degree that the orthogonality is numerically preserved, $\mathcal{D}$, can be defined by

$$\mathcal{D}(\phi) = (N-1)\frac{\sum_{n=1}^{N} |\langle \phi_n(x)\phi_n(x)\rangle|}{\sum_{n=1}^{N}\sum_{m=n+1}^{N} |\langle \phi_m(x)\phi_n(x)\rangle|} \tag{9}$$

where ϕ denotes a set of mutually orthogonal polynomials. Typically, $\mathcal{D}(T)$ is six orders of magnitude greater than $\mathcal{D}(P)$, and $\mathcal{D}(T)$ is relatively insensitive to the resolution of the domain, whereas $\mathcal{D}(P)$ increases linearly with M.

In three dimensions, the Chebyshev approximation to a function of three independent variables, $f(x, y, z)$, is given by

$$\hat{f}(x, y, z) = \sum_{i,j,k=0}^{I,J,K} C_{ijk}T_i(x')T_j(y')T_k(z') \tag{10}$$

where x', y', and z' indicate a mapping of x, y, and z to the interval $[-1, 1]$. The evaluation of the coefficients C_{ijk} follows directly from Fourier expansions,

$$C_{ijk} = \frac{1}{\|T_i\|^2\|T_j\|^2\|T_k\|^2} \times \sum_{p,q,r=1}^{P,Q,R} f(x_p'', y_q'', z_r'')T_i(x_p)T_j(y_q)T_k(z_r), \tag{11}$$

where $i, j, k = 0, \ldots, I, J, K$; $P, Q, R \geq I, J, K$; and x_p, y_q, z_r are the roots of T_P, T_Q, and T_R. The variables x_p'', y_q'' and z_r'' represent a mapping of x_p, y_q and z_r from $[-1, 1]$ to $[x_{\min}, x_{\max}]$, $[y_{\min}, y_{\max}]$, and $[z_{\min}, z_{\max}]$. The coefficients C_{ijk} minimize a least squares error norm. Combining (5), (6), (7), and (11) results in a simple expression for the coefficients C_{ijk} (Hamming, 1973).

$$C_{ijk} = \frac{1}{\|T_i\|^2\|T_j\|^2\|T_k\|^2} \sum_{p,q,r=1}^{P,Q,R} f(x_p'', y_q'', z_r'') \times \cos\left(\frac{i\pi}{P}\left(p - \frac{1}{2}\right)\right)\cos\left(\frac{j\pi}{Q}\left(q - \frac{1}{2}\right)\right) \times \cos\left(\frac{k\pi}{R}\left(r - \frac{1}{2}\right)\right) \tag{12}$$

If $f(x, y, z)$ is a continuously evaluable function, (12) is a closed-form solution for the Chebyshev coefficients. If, on the other hand, $f(x, y, z)$ is available only at discrete values of the independent variables, an interpolation method is required to obtain $f(x_p'', y_q'', z_r'')$. The interpolation routine used in the current study is a three-dimensional low-pass filter that computes a weighted sum based on the L_n norm of all the data points. Given a parametrically tabulated function, $f_t(x_t, y_t, z_t)$, $(t = 1, \ldots, T)$, an interpolated value of f at (x^*, y^*, z^*), can be calculated using

$$f(x^*, y^*, z^*) = \frac{\sum_{t=1}^{T} f_t(x_t, y_t, z_t)[(x^* - x_t)^n + (y^* - y_t)^n + (z^* - z_t)^n]^{-1}}{\sum_{t=1}^{T} [(x^* - x_t)^n + (y^* - y_t)^n + (z^* - z_t)^n]^{-1}}. \tag{13}$$

The exponent n must be a positive even number. Smaller values of n result in smoother interpolated functions. Only a cursory study of the effect of n on the goodness of fit was undertaken. A value of $n = 8$ was found to give good resolution with moderate computational effort.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

If the coefficients C_j of a Chebyshev approximation decrease rapidly with increasing j then the error associated with a Jth degree approximation is dominated by $T_{J+1}(x)$, which has equal extrema. The faster the C_j's decrease with j, the closer the Chebyshev approximation approaches a *minimax* approximation, in which the approximation is equally accurate over all regions of the domain. This is the case of a lower-order Chebyshev approximation of a monic. However, in general, a minimax approximation is not guaranteed by the choice of the basis functions, especially if the approximation minimizes a least-squares error norm.

Chebyshev approximation of nonlinear dynamical systems, including ER devices, was first carried out by Masri et al. (Ehrgott and Masri, 1992; Masri and Caughey, 1979). Masri formulated an approach in which the system nonlinearities were a function of displacement and velocity. This method is now extended to describe the nonlinearities as a function of displacement, velocity, *and* electric field.

4.2 Modeling ER Devices for Simulation and Control. In modeling and controlling ER dampers, two models are useful. In performing characterization experiments on an ER device, the motion of the device (displacement, x, and velocity, $\dot{x}$) the voltage or electric field, E, and the fluid pressure gradient or force, F, can be measured easily. Given these measurements, a function relating the force to the other three measurements,

$$\hat{F}(x, \dot{x}, E) = \sum_{i,j,k=0}^{N} C_{ijk} T_i(x') T_j(\dot{x}') T_k(E'), \qquad (14)$$

can be used to simulate the behavior of the member. If experimental results are repeatable, then a model of the form (14) describes the behavior of the device in a variety of operating conditions.

Another useful relation can be deduced from experimental measurements. Given the motion of the device and the desired force, the voltage or electric field required to produce the force can be evaluated from an expression of the form

$$\hat{E}(x, \dot{x}, |F|) = \sum_{i,j,k=0}^{N} C_{ijk} T_i(x') T_j(\dot{x}') T_k(F'). \qquad (15)$$

Equation (15) can be used to linearize the damper behavior for feedback-control applications. The unconditional dissipation of the damper requires that the force and the velocity have the same sign. So only the absolute value of the force need be specified as an independent variable. So doing, $|F|_{\min} = 0$ maps to $F' = -1$, which is on the edge of the curve-fitting domain. The absolute value of the available forces is bounded and depends on the device motion. To overcome this difficulty, if the requested force, $|F|$, is less than its minimum value, then $E = 0$ is specified, and if $|F|$ is greater than its largest possible value then a maximum E is specified. The limiting values of $|F|$ at any device motion can be evaluated from (14). In this sense, a model for control (15) requires a model for simulation (14).

4.3 Application to Test Data. Two series of tests were conducted to evaluate the predictive capabilities of the Chebyshev curve-fit. Data from the first series of tests were used to calculate the curve-fit coefficients, C_{ijk} in the approximation of Eq. (14). Data from the second series of tests were used to evaluate the ability of the curve-fit to predict the device behavior.

In the first series of tests, smooth prescribed motions, and constant electric fields were applied to the device. The electric field values were: 0.0, 0.6, 1.3, 1.9, 2.5, and 3.2 kV/mm. The accuracy of the curve-fit depends strongly on the degree, (I, J, K) of the polynomials used for curve-fitting. By evaluating the average relative root mean square error of the curve-fit, e,

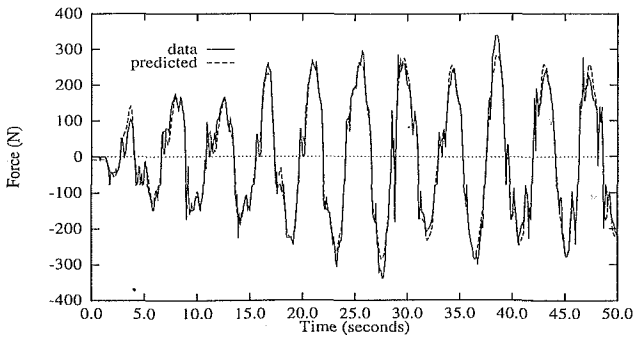


Fig. 9 Prediction of ER damper forces at E = 1.3 kV/mm, based on tests at 0.0, 0.6, 1.9, 2.4, and 3.2 kV/mm

$$e = \frac{\left[\frac{1}{T}\sum_{t=1}^{T}(F_t - \hat{F}_t)^2\right]^{1/2}}{\left[\frac{1}{T}\sum_{t=1}^{T}(F_t)^2\right]^{1/2}}, \qquad (16)$$

as a function of the polynomial degree, (I, J, K), and the number of interpolation points, (P, Q, R), an error analysis was performed for the data set. The error decreased monotonically with increasing I, J, K and P, Q, R (Gavin, 1994). Based on this sensitivity study, and the computation time required to evaluate higher degree polynomials, 50 interpolation points were used to curve-fit 11th degree polynomials, which resulted in $e = 0.115$. ($P = Q = 50$, $R = 10$ and $I = J = 11$, $K = 6$). The resulting 1008 coefficients are too numerous to list in this paper. Evaluation of Eq. (14) is very efficient and rapid as compared to Eq. (13).

In the second series of tests, data were collected at the same constant electric fields as in the first series, but the device motion was more irregular (Gavin, 1994). The curve-fit model was then verified by comparing the curve-fit force, $\hat{F}$, using the coefficients, C_{ijk} from the first set of data, and the coordinates (x, $\dot{x}$, E) of the second set of data to the measured force in the second set of data (Gavin, 1994).

As a final test of the modeling capabilities of the Chebyshev polynomial curve-fit, coefficients were calculated from the entire first data set except the data from the E = 1.3 kV/mm test. These coefficients were then used to simulate the force response from the E = 1.3 kV/mm test in the second data set. These results are shown in Fig. 9. This figure shows the ability of the curve-fit method to accurately predict the device behavior with variable electric fields.

5 Summary

As the size of the laminar passageways through which ER material flows increases, the dynamic range and response times of the ER device also increase. These conclusions have been experimentally verified using small-scale experiments. The Bingham model provides a simple and sufficiently accurate estimate of pressure gradients and force levels for design purposes.

A multivariable curve-fit results in a model which can accurately describe device-specific nonlinearities. Curve-fitting force as a function of displacement and velocity has successfully modeled a number of nonlinear elements, including ER dampers (Ehrgott and Masri, 1992; Masri and Caughey, 1979). This paper describes an extension and analysis of a method to curve-fitting in a three-dimensional domain, which includes the electric field. Curve-fitting data with Chebyshev polynomials results in good approximations to the device's behavior, which are useful for device-level simulation studies.

Some important features of ER devices are their fast response times, low-power requirements, lack of moving parts, and direct

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

transduction of an electrical command signal to a mechanical force. These devices have potential in a variety of vibration control, automotive, and robotic applications.

Acknowledgments

This material is based on work supported by the National Science Foundation under Award No. BCS-9201787 as part of the Coordinated USA Research Program on Structural Control for Safety, Performance, and Hazard Mitigation. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

The author is thankful to the referees for their helpful comments, and to the editors for coordinating the review of this paper.

References

Beyer, W. H., ed., 1984, *CRC Standard Mathematical Tables,* 27th ed., CRC Press, Boca Raton, FL.

Block, H., and Kelly, J. P., 1988, "Review Article: Electro-rheology," *Journal of Physics D: Applied Physics,* Vol. 21, pp. 1661–1677.

Brooks, D. A., 1991, "Design And Development Of Flow Based Electro-Rheological Devices," *Proceedings of the International Conference on Electro-rheological Fluids,* 15-16 Oct. 1991, Carbondale, IL, World Scientific, Singapore, pp. 367–397.

Department of Energy, 1993, "Electrorheological (ER) Fluids: A Research Needs Assessment Final Report," DOE/ER/30172, Office of Scientific and Technical Information, Oak Ridge, TN.

Ehrgott, R. C., and Masri, S. F., 1992, "Modeling the oscillatory dynamic behavior of electrorheological materials in shear," *Smart Materials and Structures,* Vol. 1, pp. 275–285.

Gamota, D. R., and Filisko, F. E., 1991, "Dynamic mechanical studies of electrorheological materials: Moderate frequencies," *Journal of Rheology,* Vol. 35, No. 3, pp. 399–425.

Gamota, D. R., and Filisko, F. E., 1991, "High frequency dynamic mechanical study of an aluminosilicate electrorheological material," *Journal of Rheology,* Vol. 35, No. 7, pp. 1411–1425.

Gavin, H. P., 1994, "Electrorheological Dampers for Structural Vibration Suppression," Ph.D. Dissertation, The University of Michigan, Ann Arbor, MI.

Gavin, H. P., Hanson, R. D., and Filisko, F. E., 1996, "Electrorheological Dampers I: Analysis and Design," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 669–675.

Jordan, T. C., and Shaw, M. T., 1989, "Electrorheology," *IEEE Transactions on Electrical Insulation,* Vol. 24, No. 5, pp. 849–878.

Masri, S. F., and Caughey, T. K., 1979, "A Nonparametric Identification Technique for Nonlinear Dynamic Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 433–447.

McClamroch, N. H., Ortiz, D. S., Gavin, H. P., and Hanson, R. D., 1994, "Electrorheological Dampers and Semi-Active Structural Control," *Proceedings of the 33rd IEEE Conference on Decision and Control,* Orlando, FL.

McClamroch, N. H., and Gavin, H. P., 1995, "Structural Control using Feedback Operated Electrorheological Dampers," *34th IEEE Conference on Decision and Control,* New Orleans, LA, submitted for publication.

Phillips, R. W., 1969, "Engineering Applications of Fluids with a Variable Yield Stress," Ph.D. Dissertation, Department of Mechanical Engineering, University of California, Berkeley, CA.

Hamming, R. W., 1986, *Numerical Methods for Scientists and Engineers,* Dover Press, New York.

Stanway, R., Sproston, J. L., and Stevens, N. G., 1985, "Non-Linear Identification of an Electrorheological Vibration Damper," *IFAC Identification and System Parameter Estimation,* pp. 195–200.

Stanway, R., Sproston, J. L., and Firoozian, 1989, "Identification of the Damping Law of an Electro-Rheological Fluid: A Sequential Filtering Approach," ASME *Journal of Dynamic Systems, Measurement, and Control,* Vol. 111, pp. 91–96.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Rotation of a Clamped Spherical Ball due to Linear Reciprocating Motion

E. S. Edelstein

J. J. Blech

Faculty of Mechanical Engineering,
Technion—Israel Institute of Technology,
Haifa 32000, Israel

A perfect spherical ball, symmetrically clamped between two identical surfaces, was found to rotate due to linear reciprocating excitation. A plausible theoretical explanation of this phenomenon follows. The rotation of the ball is a combination of the magnification of the dynamic response of the ball in the clamping system, near its natural frequencies, with frictional stick-slip at the contact areas. The ball rotates when the system is excited by a reciprocating motion not collinear with the preloading direction. The theoretical analysis of such rotation was corroborated by testing. This rotation may result in wear because of the slip motion involved. This work is focused on balls clamped between spherical (concave) surfaces. However, such rotation can also be developed in cylindrical hinges and rolling bearings, naturally loaded between a pair of surfaces. Furthermore, rotation of balls may develop in a nonrotating ball bearing under dynamic environmental conditions.

1 Introduction

A perfect spherical ball, preloaded symmetrically, by relatively high forces, between two identical surfaces (Fig 1), was found to rotate when the system was excited to linear harmonic motion. Intuitively, the symmetry of the system and the periodic linear excitation, leads to expect that there will be no development of moments on the ball, to cause its rotation.

This phenomenon was accidentally noticed, during a dynamic environmental condition test of an installation of a spherical high pressure vessel. The vessel, clamped by a force about 50 times the vessel's weight, changed its relative position to the clamp, after sinusoidal cycling having an amplitude of 4.2g was induced. No explanation of this rotation was found in the literature. Indeed, as in this case, such a rotation may pose a major obstacle in a design.

A plausible theoretical explanation of this phenomenon follows. The rotation of the ball is a combination of the magnification of the dynamic response of the ball, in the clamping system near its natural frequencies, with frictional stick-slip at the contact areas, when the system is excited by a reciprocating motion not collinear with the preloading direction. Two cases are analyzed here: In the first one, the ball is clamped between two parallel flat plates. In this case the compliance of the clamping surfaces is mainly in the normal direction. In the second case, the ball is clamped between two cavities with a spherical surfaces. The elastic compliance of the cavities, at the contact areas, exists in both directions (normal and tangential). Furthermore, in this case, the mutual interaction of the natural frequencies, in both directions, on the ball rotational response is investigated. Both cases were experimentally tested, and the test results are compared to the theoretical analysis.

Although, this article is focused on spherical balls clamped between spherical surfaces, such rotation as a response to reciprocating motion, can be developed in hinges and rolling bearings, naturally loaded between two surfaces. An interesting outcome of this research regards non rotating rolling bearings. Rotation of balls, in a non rotating ball bearing, may develop under dynamic environmental conditions. This rotation of the balls will result in wear because of the slip motion involved.

2 An Explanation of the Phenomenon

Under static conditions the ball, clamped between two surfaces, is loaded by two localized contact forces, normal to the surfaces, opposing each other, and equal in magnitude. When the system is brought into periodic linear motion, noncollinear with the preloading forces, friction forces are developed between the ball and the clamping parts, at the contact points. The line of action of those forces is tangential to the ball and its clamping surfaces (normal to the preloading forces). The excitation motion can be resolved into two components: one, in the preloading direction, and the other in the tangential direction. The first one results in relative displacement between the ball and the clamping surfaces (as a mass between two springs on a vibrating base). The second component results in friction forces which compel the ball to move in a manner coincident with the clamping surfaces.

From the point of view of an observer attached to the motion exciting device a D'Alembert "reverse effective force" is developed on the ball at its center. The component of this inertia force in the preloading direction results in a relative displacement between the ball and the clamping surfaces, in this direction. The other component of this inertia force, in the tangential direction, results in friction forces. Due to this relative displacement the normal forces vary, as one increases the other decreases. If momentarily, the moment of the inertia force with respect to one of the contact points is higher then the traction of the surface at the opposite contact point a slip displacement will take place between the ball and the clamping surface. This slip movement will occur on the side were the normal force induced by the surface on the ball is the lower one. The higher normal force at the other surface will not be able to prevent this slip motion, thus the ball rolls on this surface. Compared to slip resistance forces, rolling resisting forces are negligible.

In order to complete this discussion it has to be shown that under periodic excitation those slip displacements will result in a net rotation of the ball. Let us assume that at a certain time the excitation acceleration is to the left (see Fig. 1). An inertia force $f(t)$ is developed on the ball acting to the right. The normal force N_2 increases while N_1 decreases. If at an instant

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Dec. 22, 1993; final revision, Jan. 5, 1995. Associate Technical Editor: J. T. Jenkins.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

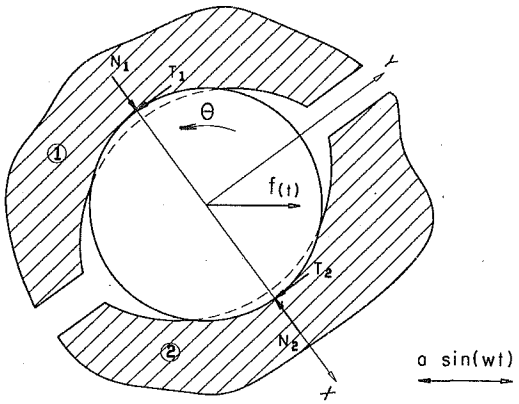


Fig. 1 The forces acting on a ball clamped between two spherical surfaces in its center position

the moment of the inertia force with respect to contact point #2 will be higher then the traction force moment of surface #1 with respect to the same point (contact point #2), a rotational displacement of the ball CW will occur. After half a period, the acceleration of the system will be to the right. An inertia force $f(t)$ is now developed on the ball to the left. The normal force N_1 increases and N_2 decreases. If at an instant the moment of the inertia force with respect to contact point #1 will be higher then the traction force moment of surface #2 with respect to the same point (contact point #1), a rotational displacement of the ball, again, CW will occur.

Finally, under a periodic external dynamic condition this slip occurs once on one surface and then on the other surface. The ball pivots on the surfaces alternately, simultaneously with the change of direction of the inertia force. Thus, a net rotational motion ensues. It has to be emphasized that near the natural frequency the displacement of the ball is magnified and even separation may occur between the ball and the clamping surfaces. In such a case, no friction coefficient will be able to prevent this rotation.

3 The Kinematic Constraints and the Equations of Motion

In the previous explanation of the phenomenon, the tangential compliance of the clamping parts was not considered. However, in the general case of a ball clamped between two spherical surfaces, the surfaces comply elastically in both directions, normal and tangential, when the ball changes its position relative to its assembled configuration. Furthermore, as the ball moves, the direction of the normal and tangential force components (in inertial space) changes (Fig. 2).

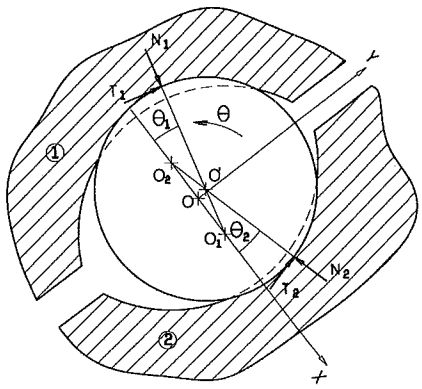


Fig. 2 The forces acting on a ball clamped between two spherical surfaces in its off center position

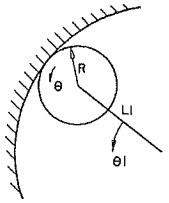


Fig. 3 The planetary mechanism representing the kinematic state

In the normal direction the magnitudes of the elastic forces depend on the distance between the ball center and the cavities' centers. The elastic forces in the tangential direction depend on the elastic displacements in that direction. In the tangential direction three different types of relative displacements are possible: *Rolling*—as the ball rolls on the cavity surface. *Slipping*—as the ball slips on the cavity surface. *Sticking (elastically)*—as the cavity surface complies elastically with the displacement of the ball. In reality, the relative displacement between the ball and the cavity is a combination of the three modes. In order to distinguish between the different types of displacements, and their contribution to the forces acting on the ball, the kinematic behavior of a planetary mechanism, shown in Fig. 3, was adapted. The kinematic constraint in such a mechanism is

$$L_1 d\theta_1 = -R\, d\theta. \tag{1}$$

Thus, in our case, the part of the ball displacement satisfied by Eq. (1) is in the rolling mode. The remaining part of the relative displacements is a towing displacement. The towing displacement is the one which results in the force acting on the ball, as the rolling forces are negligible. This towing force can appear in two ways: If it is smaller then the traction capability of the surface, then this force will be an elastic force. However, if this force is higher then the traction capability of the surface, then it will be a "slip" force.

We are now able to summarize all the forces acting on the ball.

In the normal direction:

$$N_1 = f_n\{(L_1 - L_0), \dot{L_1}\};$$
$$N_2 = f_n\{(L_2 - L_0), \dot{L_2}\}; \quad f_n\{L_0\} = 0, \tag{2}$$

where L_0 is the distance, at rest, between the ball center and the cavity center.

In the tangential direction:

for rolling $\quad T_1 = 0; \quad T_2 = 0,$

for slipping $\quad |T_1| = |N_1\, \mu_d|; \quad |T_2| = |N_2\, \mu_d|,$

For elastic displacement

$$T_1 = f_t\{[R(\theta - \theta_1^s) + L_1\theta_1], (R\dot{\theta} + L_1\dot{\theta}_1)\},$$
$$T_2 = f_t\{[R(\theta - \theta_2^s) + L_2\theta_2], (R\dot{\theta} + L_2\dot{\theta}_2)\}, \tag{3}$$

where θ is the total angular displacement of the ball, θ_1^s is the total slip angular displacement of the ball with respect to surface #1, θ_2^s is the total slip angular displacement of the ball with respect to surface #2. Denoting X_1^0, Y_1^0, and X_2^0, Y_2^0 as the coordinates of the centers of the cavities, we have.

$$L_1 = \sqrt{(x - X_1^0)^2 + (y - Y_1^0)^2}; \quad \tan\theta_1 = \frac{y - Y_1^0}{x - X_1^0}$$

$$L_2 = \sqrt{(x - X_2^0)^2 + (y - Y_2^0)^2}; \quad \tan\theta_2 = \frac{y - Y_2^0}{x - X_2^0}.$$

In order to simplify the representation of the equations of

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

motion, two additional variables are defined (for each contact point) E_{ff} and S_{ff} (E_{ff1}, E_{ff2}, S_{ff1}, S_{ff2}).

If the conditions for slip appear,
then

$$E_{ff} = 0 \quad \text{and} \quad S_{ff} = 1,$$

else

$$E_{ff} = 1 \quad \text{and} \quad S_{ff} = 0.$$

Thus

$$\theta_1^s = \int d\theta|_{s_{ff1}=1}; \quad \theta_2^s = \int d\theta|_{s_{ff2}=1}.$$

Finally, the equations of motion for harmonic excitation are

$$m\ddot{x} - f_n\{(L_1 - L_0), \dot{L}_1\}[\cos\theta_1 - \mu_d |\sin\theta_1| S_{ff1} \operatorname{sgn}(R\dot{\theta} + L_1\dot{\theta}_1)]$$
$$+ f_n\{(L_2 - L_0), \dot{L}_2\}[\cos\theta_2 - \mu_d |\sin\theta_2| S_{ff2} \operatorname{sgn}(R\dot{\theta} + L_2\dot{\theta}_2)]$$
$$+ f_t\{[R(\theta - \theta_1^s) + L_1\theta_1], (R\dot{\theta} + L_1\dot{\theta}_1)\} \sin\theta_1 E_{ff1}$$
$$- f_t\{[R(\theta - \theta_2^s) + L_2\theta_2], (R\dot{\theta} + L_2\dot{\theta}_2)\} \sin\theta_2 E_{ff2}$$
$$= m a \sin(\omega t) \cos\alpha_1 \quad (4)$$

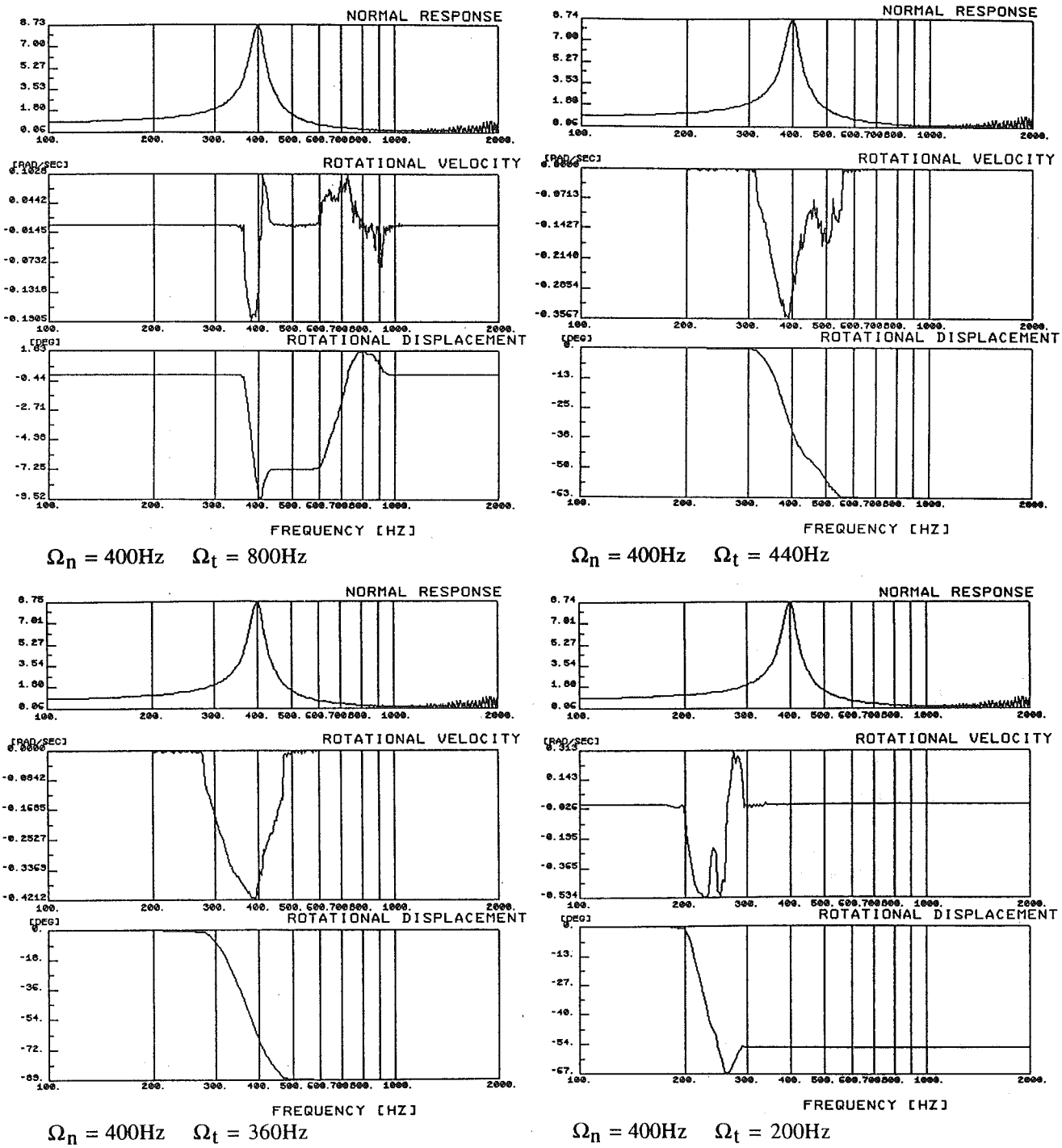


Fig. 4 The predicted rotational frequency response for several natural frequencies in the tangential direction

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

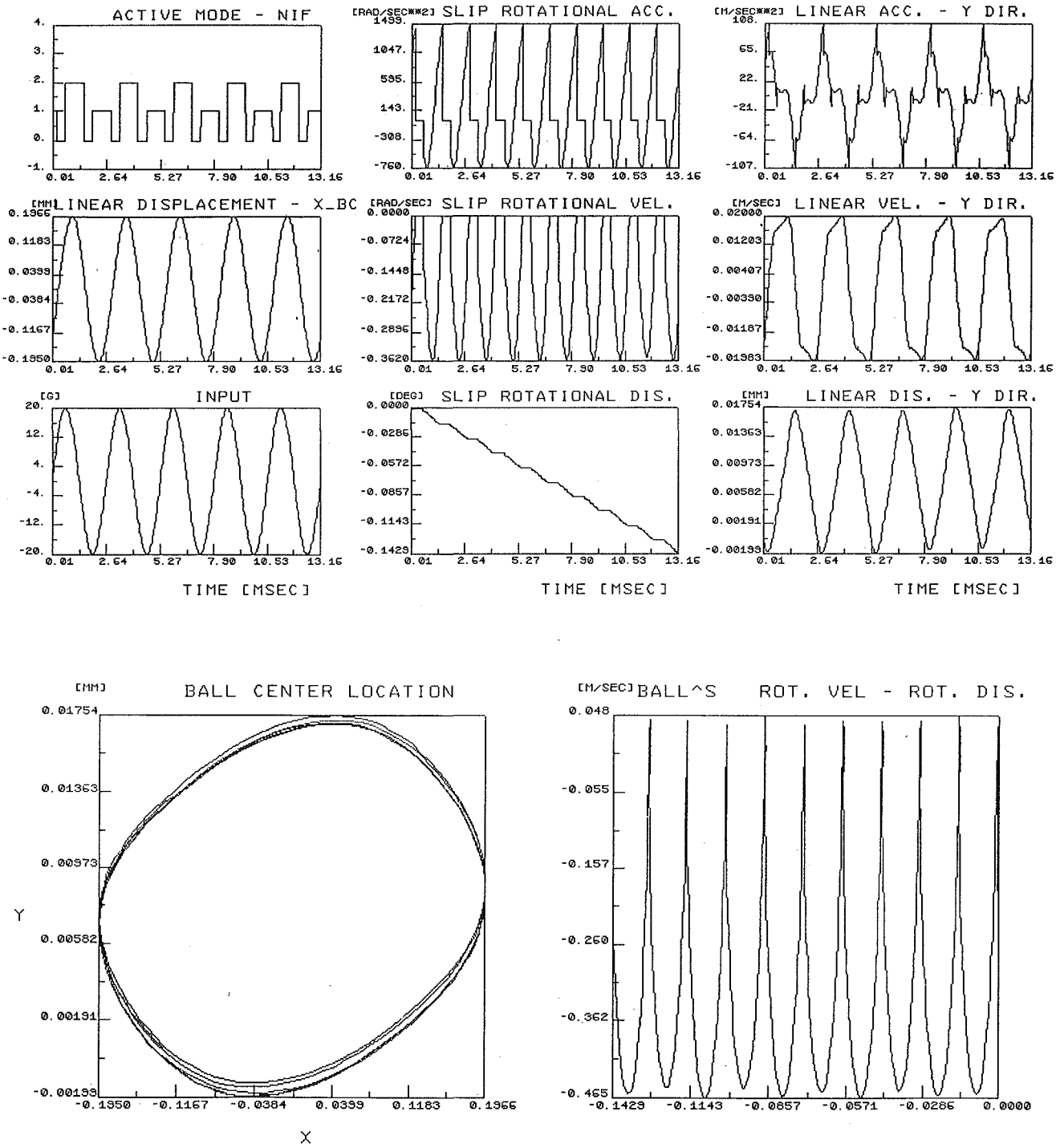


Fig. 5 The predicted rotational frequency response due to excitation at a constant frequency near the natural frequency in the normal direction

$$m\,\ddot{y} - f_n\{(L_1 - L_0), \dot{L_1}\}[\sin\theta_1 + \mu_d \cos\theta_1\, S_{ff1}\, \mathrm{sgn}(R\,\dot{\theta} + L_1\,\dot{\theta}_1)]$$
$$- f_n\{(L_2 - L_0), \dot{L_2}\}[\sin\theta_2 + \mu_d \cos\theta_2\, S_{ff2}\, \mathrm{sgn}(R\,\dot{\theta} + L_2\,\dot{\theta}_2)]$$
$$- f_t\{[R(\theta - \theta_1^s) + L_1\,\theta_1], (R\,\dot{\theta} + L_1\,\dot{\theta}_1)\}\cos\theta_1\, E_{ff1}$$
$$+ f_t\{[R(\theta - \theta_2^s) + L_2\,\theta_2], (R\,\dot{\theta} + L_2\,\dot{\theta}_2)\}\cos\theta_2\, E_{ff2} = m\,a\,\sin(\omega t)\,\sin\alpha_1 \quad (5)$$

$$J\,\ddot{\theta} + \{f_n\{(L_1 - L_0), \dot{L_1}\}\,\mu_d\, S_{ff1}\,\mathrm{sgn}(R\,\dot{\theta} + L_1\,\dot{\theta}_1)$$
$$+ f_n\{(L_2 - L_0), \dot{L_2}\}\,\mu_d\, S_{ff2}\,\mathrm{sgn}(R\,\dot{\theta} + L_2\,\dot{\theta}_2)$$
$$+ f_t\{[R(\theta - \theta_1^s) + L_1\,\theta_1], (R\,\dot{\theta} + L_1\,\dot{\theta}_1)\}E_{ff1}$$
$$+ f_t\{[R(\theta - \theta_2^s) + L_2\,\theta_2], (R\,\dot{\theta} + L_2\,\dot{\theta}_2)\}E_{ff2}\}R = 0 \quad (6)$$

Under the following conditions slip occurs between the ball and surface #1.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

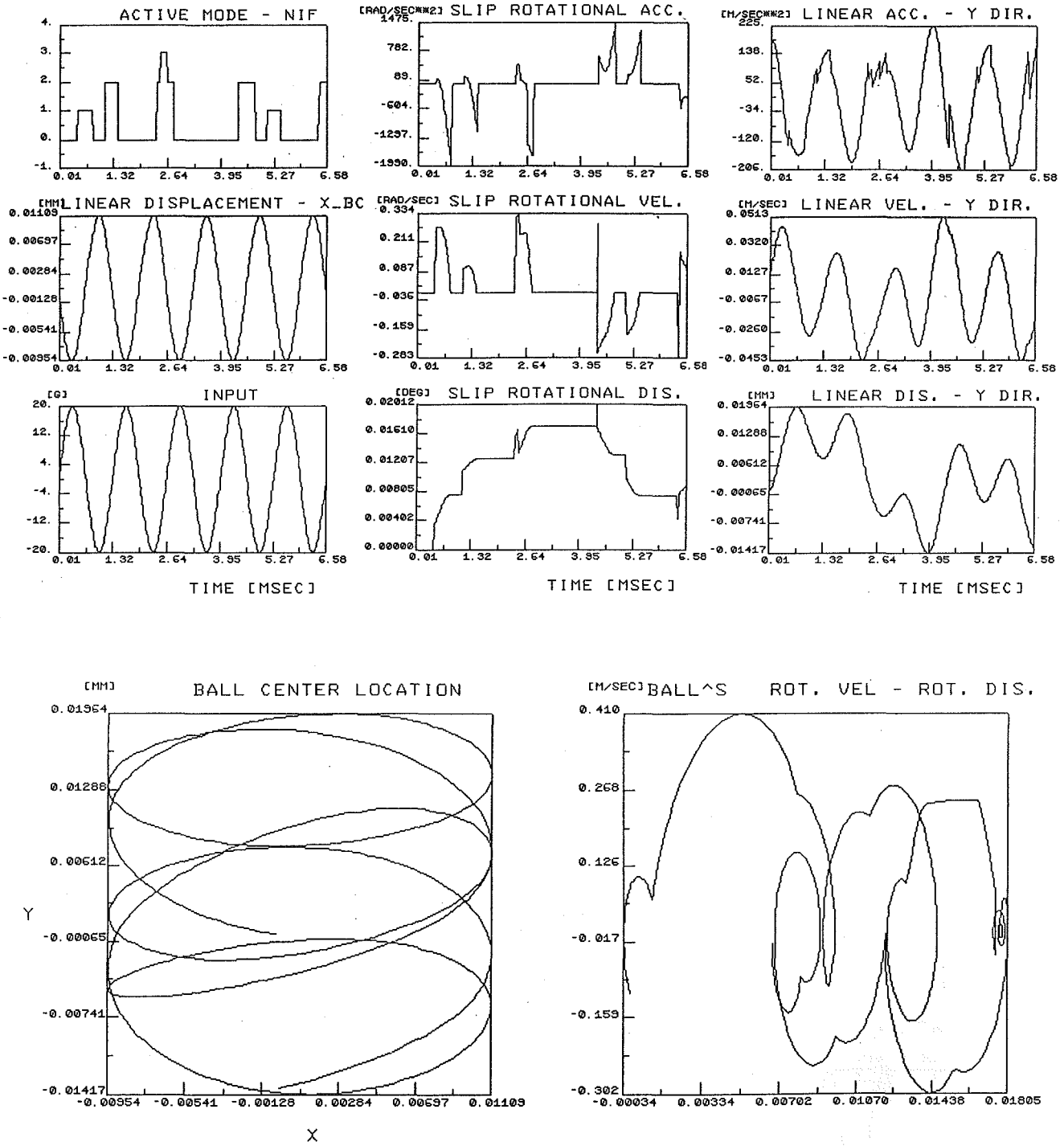


Fig. 6 The predicted rotational frequency response due to excitation at a constant frequency near the natural frequency in the tangential direction

If

$$S_{ff1} = 1 \tag{7}$$

or

$$|f_t\{[R\,(\theta - \theta_1^s) + L_1\,\theta_1], (R\,\dot{\theta} + L_1\,\dot{\theta}_1)\}| \geq |f_n\{(L_1 - L_0), \dot{L}_1\}|\,\mu_s$$

then

$$S_{ff1} = 1 \qquad E_{ff1} = 0.$$

There is no slip between the ball and surface #1 (elastic compliance of surface #1) if

$$(R\,\dot{\theta} + L_1\,\dot{\theta}_1 = 0 \quad \text{or} \quad E_{ff1} = 1) \tag{8}$$

and

$$|f_t\{[R\,(\theta - \theta_1^s) + L_1\,\theta_1], (R\,\dot{\theta} + L_1\,\dot{\theta}_1)\}| < |f_n\{(L_1 - L_0), \dot{L}_1\}|\,\mu_s$$

then

$$S_{ff1} = 0 \qquad E_{ff1} = 1.$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

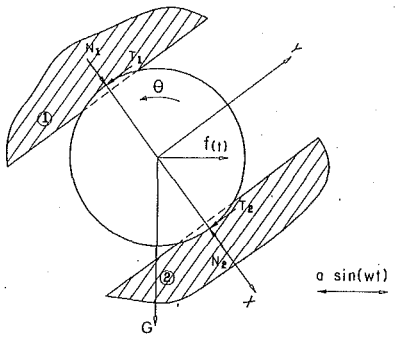


Fig. 7 The forces acting on a ball clamped between two flat plates

Slip occurs between the ball and surface #2
if

$$S_{ff2} = 1 \tag{9}$$

or

$$|f_t\{[R\,(\theta - \theta_2^s) + L_2\,\theta_2],\,(R\,\dot{\theta} + L_2\,\dot{\theta}_2)\}| \geq |f_n\{(L_2 - L_0),\,\dot{L}_2\}|\,\mu_s$$

then

$$S_{ff2} = 1 \qquad E_{ff2} = 0.$$

There is no slip between the ball and surface #2 (elastic compliance of surface #2)
if

$$(R\,\dot{\theta} + L_2\,\dot{\theta}_2 = 0 \quad \text{or} \quad E_{ff2} = 1) \tag{10}$$

and

$$|f_t\{[R\,(\theta - \theta_2^s) + L_2\,\theta_2],\,(R\,\dot{\theta} + L_2\,\dot{\theta}_2)\}| < |f_n\{(L_2 - L_0),\,\dot{L}_2\}|\,\mu_s$$

then

$$S_{ff2} = 0 \qquad E_{ff2} = 1.$$

4 The Rotational Frequency Response for a Linear Compliance

The rotational frequency response is a function of a variety of parameters. Moreover, part of those parameters are not constant, as for instance friction coefficients. However, the natural frequencies of the system are the main parameters that govern the rotational frequency response. For this analysis the compliances are taken as linear and the damping as viscous. Thus

$$f_n\{(L_1 - L_0),\,\dot{L}_1\} = K_n\,(L_1 - L_0) + C_n\,\dot{L}_1$$

$$f_n\{(L_2 - L_0),\,\dot{L}_2\} = K_n\,(L_2 - L_0) + C_n\,\dot{L}_2$$

$$f_t\{[R\,(\theta - \theta_1^s) + L_1\,\theta_1],\,(R\,\dot{\theta} + L_1\,\dot{\theta}_1)\} = K_t\,[R\,(\theta - \theta_1^s) + L_1\,\theta_1] + C_t\,(R\,\dot{\theta} + L_1\,\dot{\theta}_1)$$

$$f_t\{[R\,(\theta - \theta_2^s) + L_2\,\theta_2],\,(R\,\dot{\theta} + L_2\,\dot{\theta}_2)\} = K_t\,[R\,(\theta - \theta_2^s) + L_2\,\theta_2] + C_t\,(R\,\dot{\theta} + L_2\,\dot{\theta}_2)$$

where K_n and K_t are the elastic constants of one surface in the normal and tangential directions respectively, C_n and C_t are the damping coefficients of the surfaces as a system in the normal and tangential directions, respectively.

If the tangential compliance is negligible, (one natural frequency "rules," in the normal direction only) the ball rotates in one direction due to excitation below the natural frequency and reverses its rotation when excited above the natural frequency. This phenomenon is due to approximately 180 deg phase reversal when operating below and above the natural frequency. Consequently, the pivoting of the ball on the clamping parts above the natural frequency is opposite to the pivoting beneath the natural frequency. However, the force normal to the preloading direction is coincident with the exciting motion.

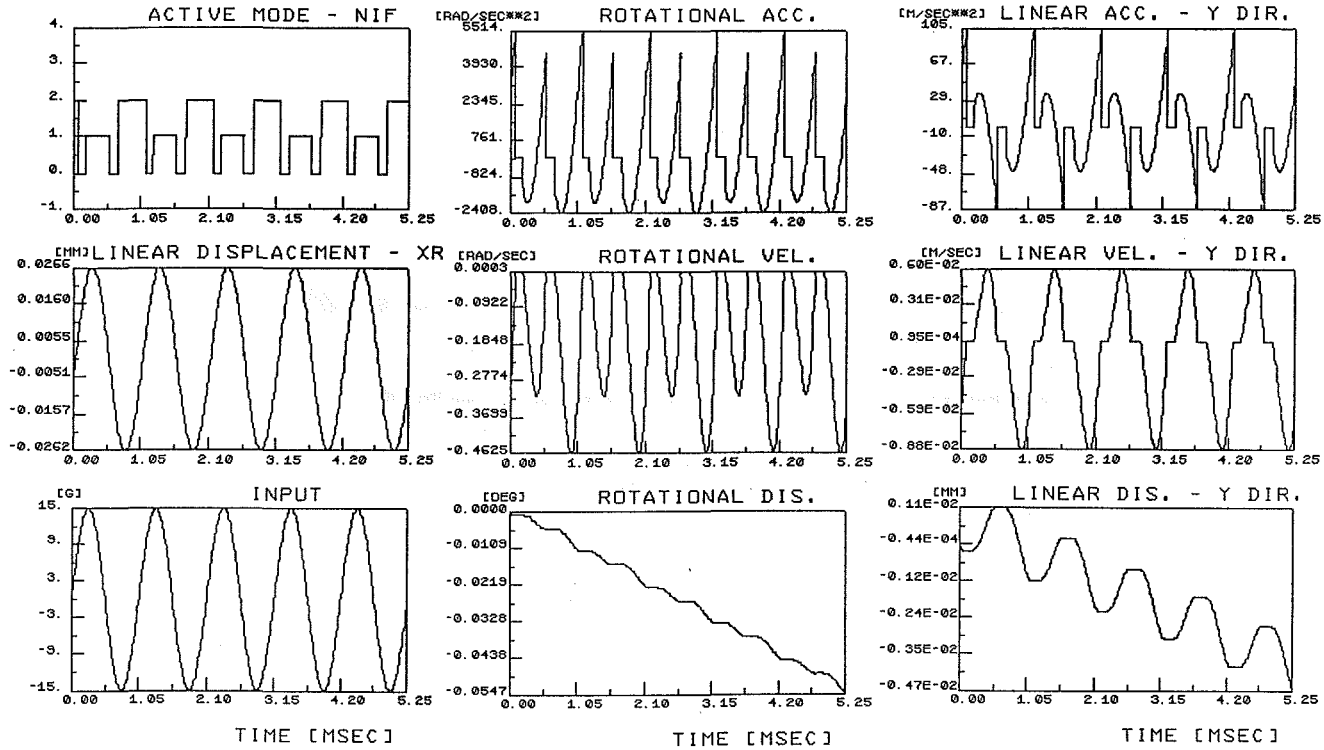


Fig. 8 The predicted response of a ball clamped between two parallel planes to harmonic excitation

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

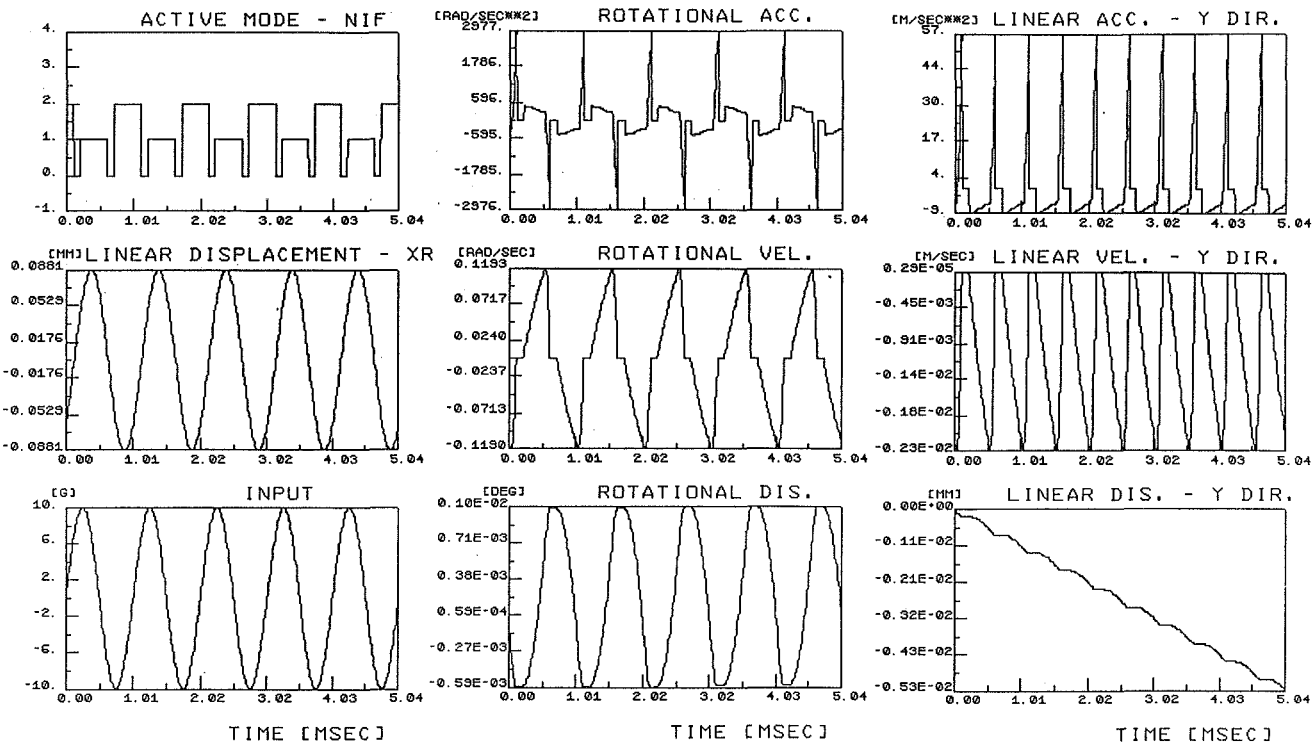


Fig. 9 The predicted response of a ball clamped between two vertical parallel planes to horizontal harmonic excitation

The result is a reversal rotational response to excitation at a frequency above the natural frequency.

As the compliance in the tangential direction increases, the reversal in rotation diminishes. Consequently, when the compliance of the clamping parts results in natural frequencies which are close to each other, the ball rotates in one direction only. Figure 4 shows the rotational frequency response in four cases. In these cases the natural frequency in the normal direction (Ω_n) is 400 Hz and the natural frequencies in the tangential direction (Ω_t) are 800 Hz, 440 Hz, 360 Hz, and 200 Hz. It can be seen in Fig. 4 that part of the curve of the rotational velocity is smooth and part of it is not. Thus, the rotational frequency response due to excitation near the natural frequency in the normal direction is periodic. This can be seen in Fig. 5. On the other hand, the rotational frequency response due to excitation near the natural frequency in the tangential direction is very irregular. This can be seen in Fig. 6.

5 Ball Clamped Between Parallel Flat Plates

It is assumed that the compliance of each surface is in the normal direction only. In order to simplify the analysis the following additional assumptions are made:

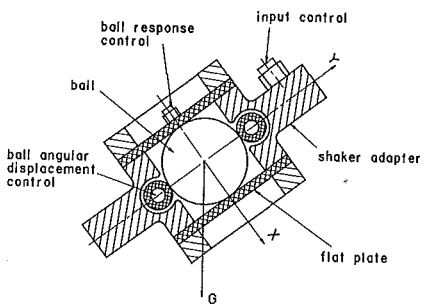


Fig. 10 Test apparatus—ball clamped between two parallel plates

1 All characteristics of both surfaces are identical.
2 The compliance of each surface is linear.
3 The damping of the surfaces compliance is viscous.
4 The friction coefficient has two levels: static friction and dynamic friction.

Under those assumptions the equations of motion, in a coordinate system attached to the clamping surfaces are

$$m\ddot{x} + C_n\dot{x} + 2K_n x = -ma\sin(\omega t)\cos\alpha_1 + mg\cos\alpha_2, \quad (12)$$

$$m\ddot{y} = T_1 - T_2 - mg\sin\alpha_2 - ma\sin(\omega t)\sin\alpha_1, \quad (13)$$

$$J\ddot{\theta} = -(T_1 + T_2)R. \quad (14)$$

The equation in the x direction is decoupled from the other equations of the system, and its solution is

$$x = X\sin(\omega t - \phi_n) + \frac{g\cos\alpha_2}{\Omega_n^2} \quad (15)$$

where

$$X = \frac{a\cos\alpha_1/\omega^2}{\sqrt{[1-(\omega/\Omega_n)^2]^2 + [2\eta_n(\omega/\Omega_n)]^2}};$$

$$\tan\phi_n = \frac{2\eta_n\omega/\Omega_n}{1-(\omega/\Omega_n)^2}$$

$$\Omega_n = \sqrt{\frac{2K_n}{m}}; \quad \eta_n = \frac{C_n}{2m\Omega_n}. \quad (16)$$

T_1 and T_2 are the tangential forces. Their magnitude is a function of the dynamic conditions. Four different modes of response are possible:

1 No slip occurs. In this mode,

$$|T_1| < |N_1\mu_s| \quad \text{and} \quad |T_2| < |N_2\mu_s|. \quad (17)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

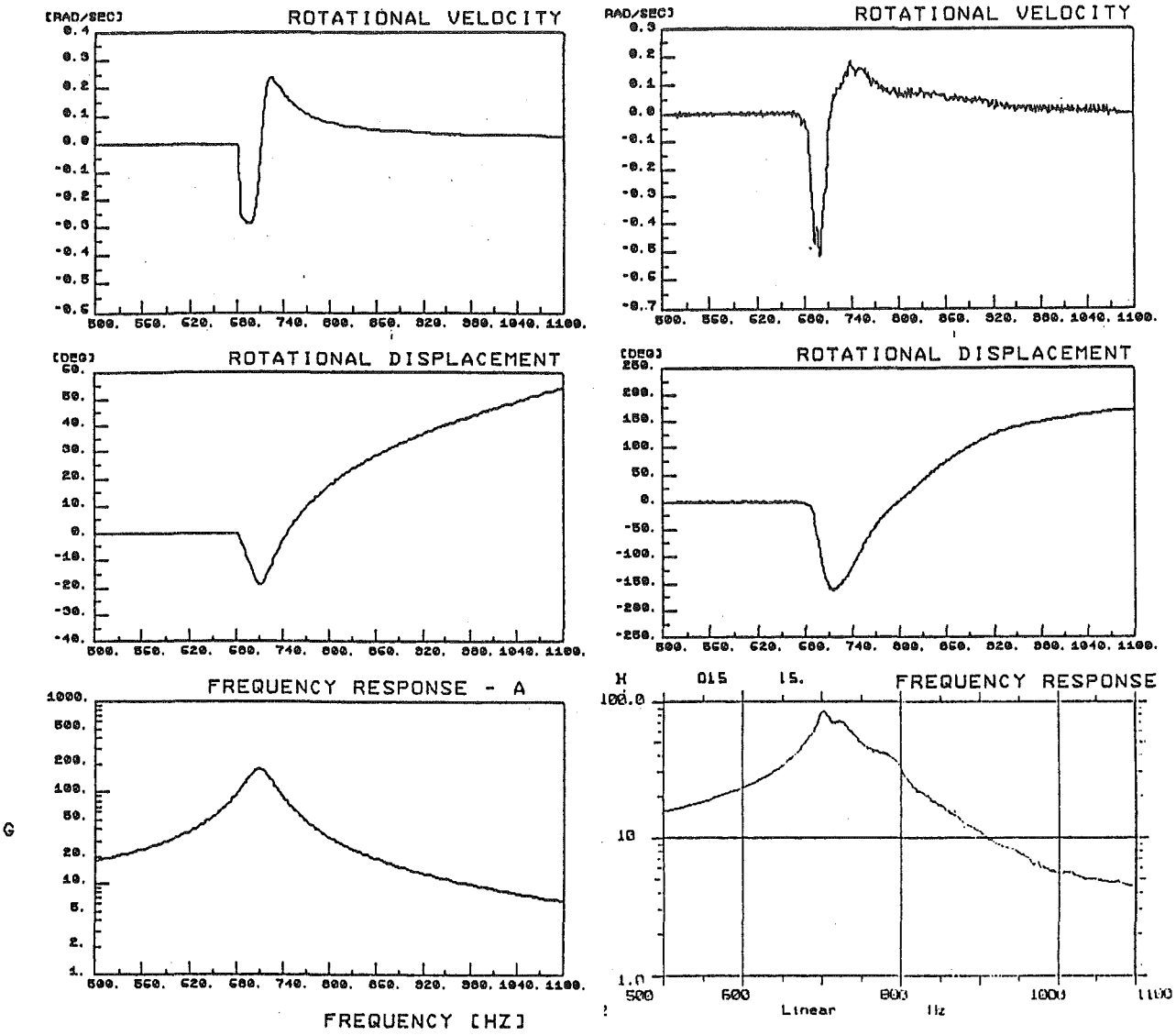


Fig. 11 Test and simulation results—ball clamped between flat plates

Thus

$$-T_1 = T_2 = \frac{m\ a\ \sin\alpha_1\ \sin(\omega t) + m\ g\ \sin\alpha_2}{2}$$

and

$$\ddot{\theta} = 0 \quad \ddot{y} = 0.$$

2 Slip occurs between the ball and surface #1. In this mode.

$$|T_1| \geq |N_1\ \mu_s| \quad \text{and} \quad |T_2| < |N_2\ \mu_s| \tag{18}$$

or

$$R\ \dot{\theta} - \dot{y} \neq 0 \quad \text{and} \quad R\ \dot{\theta} + \dot{y} = 0.$$

Thus

$$T_1 = |N_1\ \mu_d|\ \text{sgn}[\dot{y} - R\ \dot{\theta}]$$

$$\ddot{\theta} = \{m\ a\ \sin\alpha_1\ \sin(\omega t) + m\ g\ \sin\alpha_2 - 2\ N_1\ \mu_d\ \text{sgn}[\dot{y} - R\ \dot{\theta}]\}\ \frac{R}{J + m\ R^2}$$

and

$$\ddot{y} = -\ R\ \ddot{\theta}; \quad T_2 = -\frac{J\ \ddot{\theta} + T_1\ R}{R}.$$

3 Slip occurs between the ball and surface #2. In this mode

$$|T_1| < |N_1\ \mu_s| \quad \text{and} \quad |T_2| \geq |N_2\ \mu_s| \tag{19}$$

or

$$R\ \dot{\theta} - \dot{y} = 0 \quad \text{and} \quad R\ \dot{\theta} + \dot{y} \neq 0.$$

Thus

$$T_2 = |N_2\ \mu_d|\ \text{sgn}[\dot{y} + R\ \dot{\theta}]$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\ddot{\theta} = \{m\,a\,\sin\alpha_1\,\sin(\omega t) + m\,g\,\sin\alpha_2 - 2\,N_2\,\mu_d\,\mathrm{sgn}[\dot{y} + R\,\dot{\theta}]\}\,\frac{R}{J + m\,R^2}$$

and

$$\ddot{y} = R\,\ddot{\theta};\quad T_1 = -\frac{J\,\ddot{\theta} + T_2\,R}{R}.$$

4 Slip occurs between the ball and both surfaces. In this mode

$$R\,\dot{\theta} - \dot{y} \neq 0 \quad \text{and} \quad R\,\dot{\theta} + \dot{y} \neq 0 \tag{20}$$

or

$$R\,\dot{\theta} - \dot{y} \neq 0 \quad \text{and} \quad |T_2| < |N_2\,\mu_s|$$

or

$$R\,\dot{\theta} + \dot{y} \neq 0 \quad \text{and} \quad |T_1| < |N_1\,\mu_s|.$$

Thus

$$T_1 = -N_1\,\mu_d\,\mathrm{sgn}[\dot{y} - R\,\dot{\theta}];\quad T_2 = -N_2\,\mu_d\,\mathrm{sgn}[\dot{y} + R\,\dot{\theta}]$$

and

$$\ddot{y} = \frac{-N_1\,\mu_d\,\mathrm{sgn}[\dot{y} - R\,\dot{\theta}] + N_2\,\mu_d\,\mathrm{sgn}[\dot{y} + R\,\dot{\theta}] - m\,g\,\sin\alpha_2 - m\,a\,\sin(\omega t)\,\sin\alpha_2}{m}$$

$$\ddot{\theta} = \{N_1\,\mu_d\,\mathrm{sgn}[\dot{y} - R\,\dot{\theta}] + N_2\,\mu_d\,\mathrm{sgn}[\dot{y} + R\,\dot{\theta}]\}\,\frac{R}{J}.$$

Equations (13), (14) were numerically integrated subject to conditions (17)–(20). The dynamic response of the ball to harmonic excitation is shown in Fig. 8.

The horizontal lines in the rotational direction and in the y direction graphs in Fig. 8 represent time intervals where no displacement occurs. The following vertical lines represent a sharp transition due to the change from static to dynamic friction. The curved portions of the graphs represent the variation of the acceleration until the ball stops, which is designated by a vertical line. The rotational velocity is in one direction, thus, giving a net rotational displacement, as shown. However, in the y direction the velocity is once in one direction and once in the opposite direction. Consequently, the ball moves forwards and backwards. The total small net displacement in the y direction is as a result of gravity.

6 Numerical and Experimental Tests of the Theoretical Model

The theoretical model suggested was tested in three ways: First, in the intuitive way, where the dynamic response can be foreseen: A ball clamped between two vertical flat plates exited to horizontal reciprocating motion. In this case, the ball is attached to one surface and rolls on it down while slip occurs on the opposite surface. This results in an increasing displacement in the y direction, and a forward and backward displacement in the rotational direction (Fig. 9).

In the second test, a comparison of results was done between the general theoretical model and the unidirectional compliance model, i.e., a ball clamped between flat plates. For the two theoretical models, two different kinematic constraints were adapted. By letting the cavities radius be infinite and their compliance in the tangential direction be zero, the general case is reduced to a ball clamped between flat plates. The comparison of results of both cases for the same data gave identical results.

In the third test, an experimental investigation was performed for the ball clamped between two flat plates (Fig. 10). In this case the compliance of the clamping parts is mainly in the normal direction. A 38.1 mm diameter steel ball was clamped between two 3.7 mm "perspex" flat plates fixed. The plates were fixed at their circumferences (39 mm diameter). The unassembled distance between the two plates was smaller then the diameter of the ball, so the ball is preloaded when the system is assembled. The experiments were subsequently simulated by numerical integration of the equations of motion. All the parameters needed for the simulations (compliance, friction coefficients, natural frequencies, and damping coefficient) were experimentally determined in the assembled testing systems.

In one of the tests the ball was preloaded by a force 150 times the ball weight. This system was scanned in the 500–1100 Hz frequency range to horizontal harmonic motion with an amplitude of 15 g, at an angle of 37 deg to the preload direction. (the natural frequency of the system, ball—flat plates, is 700 Hz approximately). The input was controlled by two accelerometers, one attached to the shaker, and the second, attached to the shaker adapter, controlling the input in the normal direction.

The ball frequency response was measured by an accelerometer attached to one of the plates. The ball angular displacement was measured by a potentiometer. Subsequently, this test was numerically simulated (Fig. 11). It is assumed that the flat plates are linear springs in the normal direction. The compliance in the tangential direction is negligible.

Experiments with ball clamped between spherical cavities in thick plates (whereby the normal and tangential compliance of the surfaces is not linear) excited to linear reciprocating motion were also conducted. Test results and their analysis will appear elsewhere, due to the volume of the material involved.

7 Conclusions

A perfect spherical ball, symmetrically clamped between two identical surfaces, rotates due to linear reciprocating motion. In a system with one natural frequency, in the normal direction only, the ball rotates in one direction due to excitation below the natural frequency and reverses its rotation when excited above the natural frequency. However, as the compliance in the tangential direction increases, the reversal in rotation diminishes. Consequently, when the compliance of the clamping parts results in natural frequencies which are close to each other, the ball rotates in one direction only. Although this article is focused on spherical balls clamped between spherical surfaces, such rotation, as a response to reciprocating motion, can be developed in hinges and rolling bearings, naturally loaded between two surfaces. An interesting outcome of this research regards nonrotating rolling bearings. Rotation of balls, in a nonrotating ball bearing, may develop under dynamic environmental conditions. This rotation of the balls will result in wear because of the slip motion involved in.

This paper deals with the case of linear compliance at the contact areas. However, in the case of axisymmetrical parts clamped between thick parts it is necessary to calculate the compliance of the part according to Hertz's theory in the normal direction and Mindlin's solution in the tangential direction. This case is the subject of an upcoming paper (Edelstein and Blech, 1993).

References

Edelstein, E. S., and Blech, J. J., 1993, "Rotation of a Clamped Parts Due to Reciprocating Motion," ASME JOURNAL OF APPLIED MECHANICS, submitted for publication.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Application of Objective Rates in Mechanical Modeling of Solids

W. D. Reinhardt

R. N. Dubey
Professor,
Mem. ASME.

Department of Mechanical Engineering,
University of Waterloo,
Waterloo, Ont. N2L 3G1, Canada

A unified formulation is developed for deformation-related spins, and for objective rates based on them. The approach generalizes the underlying concepts, and allows new rates to be constructed. Mathematical and thermodynamical restrictions on these are shown. As a result, it can be demonstrated that the Eulerian strain rate is an objective rate of logarithmic strain, based on a spin easily derivable from the general form. Interrelations between other known spins and objective rates emerge very clearly. Consequences of the proposed formalism are explored in hypoelastic and in rigid-plastic constitutive relations, the latter involving purely isotropic and purely kinematic hardening. The application of the resulting models to the simple shear deformation is shown.

Introduction

In describing the constitutive behavior of a material it is often advantageous to employ an incremental, or rate, formulation. If the problem involves large deformations, it becomes necessary to define the rate of stress or strain with respect to a rotating frame of reference which renders the constitutive description independent of rigid-body motions. The goal of the present work is to show a generalization of the class of deformation-related objective rates which includes many of the rates that have been used in prior publications. The generalization leads to a simple proof that the Eulerian strain rate tensor is an objective rate of logarithmic strain.

In the following, tensors will be denoted by boldface letters. Let the finite deformation of a body be described by the deformation gradient $\mathbf{F}$. This tensor can be decomposed into a symmetric stretch tensor $\mathbf{U}$ or $\mathbf{V}$ and an orthogonal rotation tensor $\mathbf{R}$

$$\mathbf{F} = \mathbf{VR} = \mathbf{RU}. \tag{1}$$

Hence, an orthogonal transformation leads from the left stretch tensor $\mathbf{V}$ to the right stretch $\mathbf{U}$ and vice versa. The two can be interpreted as components of the same tensor in two frames which are related by a pure rotation $\mathbf{R}$. The $\mathbf{V}$-components are based on the fixed x_i-axes, and the components $\mathbf{U}$ on the z_i-axes which rotate relative to the fixed axes. The spin of the z_i-axes, or body spin, is given by

$$\boldsymbol{\Omega} = \dot{\mathbf{R}}\mathbf{R}^T, \tag{2}$$

where the superscript T denotes the transpose.

The principal components of the stretch $\mathbf{V}$ (assumed to be distinct) are denoted by λ_i, and the spin of the principal axes is $\boldsymbol{\Omega}^E$. The stretch components serve to define the logarithmic strain tensor which we will use as the measure of deformation:

$$\mathbf{e} = \ln \mathbf{V}. \tag{3}$$

Finally, the Eulerian strain rate tensor $\mathbf{D}$ and the material spin $\mathbf{W}$ can be derived as the symmetric and skew-symmetric part of the velocity gradient.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Aug. 1, 1994; final revision, Mar. 1995. Associate Technical Editor: J. T. Jenkins.

Spins and Objective Rates

Time derivatives of tensors (tensor rates) used in constitutive relations need to be objective. That is, the rate must be taken on a frame of reference which is body-fixed when the body undergoes a rigid rotation. We denote a co-rotational objective rate of a tensor $\mathbf{A}$ by

$$\dot{\mathbf{A}}^o = \dot{\mathbf{A}} + \mathbf{A}\boldsymbol{\Omega}^o_{\text{rel}} - \boldsymbol{\Omega}^o_{\text{rel}}\mathbf{A}, \tag{4}$$

where $\dot{\mathbf{A}}$ denotes the material rate with respect to the basis of $\mathbf{A}$. Therefore, the objective rate of $\mathbf{A}$, $\dot{\mathbf{A}}^o$, is the material rate associated with a rotating frame which has a spin $\boldsymbol{\Omega}^o_{\text{rel}}$ relative to the basis of $\mathbf{A}$. Unfortunately, the choice of a rotating frame satisfying the above requirement is not uniquely determined. Many examples of frames which are suitable as a basis for an objective rate have been published. Some of these rates are derived only from deformation-related quantities (e.g., Truesdell and Noll, 1965; Dienes, 1979; Dubey, 1987; Metzger and Dubey, 1987), others include stress-related parameters as well (e.g., Lee et al., 1983; Dafalias, 1983; Zbib and Aifantis, 1987; Bammann and Aifantis, 1987). However, the present discussion will be restricted to the first type.

Well-known examples of deformation-related spin tensors $\boldsymbol{\Omega}^o$ are the body spin $\boldsymbol{\Omega}$, the material spin, $\mathbf{W}$, and the spin of the principal axes of $\mathbf{V}$, $\boldsymbol{\Omega}^E$. Linear combinations of these have been considered as well, e.g., by Halleux and Donea (1986).

All these spins can be represented in a unified form if we transform to the principal axes of stretch. The spins relative to the principal frame are then obtained as

$$\Omega^{o(e)}_{ij\text{rel}} = \Omega^{o(e)}_{ij} - \Omega^{E(e)}_{ij} = q^o_{ij} D^{(e)}_{ij} \quad (i \neq j, \text{ no sum}). \tag{5}$$

In this equation, $\mathbf{D}^{(e)}$ represents the components of the Eulerian strain rate tensor $\mathbf{D}$ on the principal axes of strain. The only difference between the spins lies in the factor q^o which depends on the principal values of $\mathbf{V}$, or, equivalently, on the principal values of logarithmic strain, E_i. For $\mathbf{W}$, q^o becomes

$$q^J_{ij} = \frac{1}{\tanh (E_i - E_j)}, \tag{6}$$

for $\boldsymbol{\Omega}$ it is

$$q^z_{ij} = \frac{1}{\sinh (E_i - E_j)}, \tag{7}$$

and for the principal axes spin $\boldsymbol{\Omega}^E$

$$q^E_{ij} = 0, \tag{8}$$

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

because the spin of the principal axes relative to themselves is zero. Equations (6) and (7) are due to Mehrabadi and Nemat-Nasser (1987). Since the factor q depends on the difference of two principal strains, the spins are determined by the deviatoric part of the strain tensor. The structure of these spins suggests the following generalization:

$$q_{ij}^{o} = f(E_i - E_j) \tag{9}$$

where f must be an odd function. We assume that the principal strains are either distinct, or that a suitable value can be constructed to ensure a continuous development of the spin.

To explore the application of objective rates of strain, we let $\mathbf{A}$ in (4) equal the logarithmic strain tensor $\mathbf{e}$ and transform to the principal axes of $\mathbf{e}$ to obtain, in view of (4) and (5),

$$\left.\begin{aligned} \dot{e}_{ij}^{o(e)} = q_{ij}^{o}(E_i - E_j)D_{ij}^{(e)} \equiv p_{ij}^{o}D_{ij}^{(e)} \text{ for } i \neq j \\ \dot{e}_{ii}^{o(e)} = D_{ii}^{(e)} \end{aligned}\right\} \text{(no sum)}. \tag{10}$$

For the Zaremba-Jaumann-Noll rate (J-rate) which uses the spin $\mathbf{W}$,

$$p_{ij}^{J} = \frac{E_i - E_j}{\tanh\,(E_i - E_j)}. \tag{11}$$

The Green-McInnis rate (Green and McInnis, 1967), or Z-rate, is based on the body spin $\boldsymbol{\Omega}$ and gives

$$p_{ij}^{Z} = \frac{E_i - E_j}{\sinh\,(E_i - E_j)}. \tag{12}$$

For the Eulerian rate (E-rate) we obtain

$$p_{ij}^{E} = 0. \tag{13}$$

Obviously, the requirement

$$p_{ij}^{D} = 1 \tag{14}$$

leads to $\dot{\mathbf{e}}^{D} = \mathbf{D}$ which we call the D-rate of strain. The q^o of the corresponding spin are found to be

$$q_{ij}^{D} = \frac{1}{E_i - E_j} \tag{15}$$

which is the correct format for a generalized spin as defined by (9). The spin itself becomes, in its fixed axes representation

$$\boldsymbol{\Omega}^{D} = \boldsymbol{\Omega}^{E} + (\mathbf{Q}^{D})^{1/2}(\mathbf{D} - \dot{\mathbf{e}}^{E})(\mathbf{Q}^{D})^{1/2}, \tag{16}$$

where the principal components of $\mathbf{Q}^{D}$ are given by

$$\mathbf{Q}^{D^{(e)}} = \begin{pmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & Q_3 \end{pmatrix},$$

$$Q_k = \frac{q_{ik}^{D}q_{jk}^{D}}{q_{ij}^{D}};\ \text{(no sum)};\ i, j, k = 1 \ldots 3;\ i \neq j \neq k. \tag{17}$$

Note that the other spins have a similar tensor representation. The general form of the objective rate follows from (10) and (16) as

$$\dot{\mathbf{e}}^{o} = \dot{\mathbf{e}}^{E} + (\mathbf{P}^{o})^{1/2}(\mathbf{D} - \dot{\mathbf{e}}^{E})(\mathbf{P}^{o})^{1/2}, \tag{18}$$

where $\mathbf{P}^{o}$ is defined analogously to (17):

$$\mathbf{P}^{o^{(e)}} = \begin{pmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ 0 & 0 & P_3 \end{pmatrix},$$

$$P_k = \frac{p_{ik}^{o}p_{jk}^{o}}{p_{ij}^{o}};\ \text{(no sum)};\quad i, j, k = 1 \ldots 3;\ i \neq j \neq k. \tag{19}$$

We have seen how the concept of an objective rate of the form (10) leads to the proof that the Eulerian strain rate is an objective rate of logarithmic strain. The spin corresponding to this rate is given by (16). It seems that these facts have not been realized to date, and the paper by Hill (1978) even contains a statement that $\mathbf{D}$ "is not a direct flux" of any strain measure. The method which is presented here not only reveals that $\mathbf{D}$ is actually a rate of logarithmic strain, it also makes the derivation especially straightforward. Some consequences of these results will be explored in the following sections. Note that the D-spin is not the spin of the principal axes of $\mathbf{D}$.

The properties of the functions determining the factor q^o allow one to establish the ordering $q_{ij}^{E} \leq q_{ij}^{Z} \leq q_{ij}^{D} \leq q_{ij}^{J}$ for the elements with $E_i > E_j$. The opposite ordering holds for $E_i < E_j$. Since all spins depend on the components of $\mathbf{D}^{(e)}$ identically, the same kind of ordering holds for the relative spins on the principal axes of strain

$$\Omega_{ij\text{rel}}^{E(e)}\ (\equiv 0) \leq \Omega_{ij\text{rel}}^{(e)} \leq \Omega_{ij\text{rel}}^{D(e)} \leq W_{ij\text{rel}}^{(e)},\quad i \neq j. \tag{20}$$

If the spins are taken with respect to any other frame, the relative spins (20) are subjected to an orthogonal transformation, and a spin tensor is added which is the same for all spins. In general, the orthogonal transformation may change the ordering for individual components of the spin tensor. However, if the rotations and spins are strictly two-dimensional, the above ordering holds for any frame of reference.

A similar property can be derived for the objective rates of strain based on these spins. In fact, since $f(E_i - E_j)$ has the same sign as $E_i - E_j$, the ordering

$$p_{ij}^{E} \leq p_{ij}^{Z} \leq p_{ij}^{D} \leq p_{ij}^{J} \tag{21}$$

holds for all i, j $(i \neq j)$. The diagonal elements of $\dot{\mathbf{e}}^{o}$ are identical for all objective rates, so that for each component on the principal axes follows

$$\dot{e}_{ij}^{E(e)} \leq \dot{e}_{ij}^{Z(e)} \leq \dot{e}_{ij}^{D(e)} \leq \dot{e}_{ij}^{J(e)}. \tag{22}$$

In a later section of this paper the significance of these results in constitutive modelling will be demonstrated.

Constitutive Models

Large deformation hypoelastic constitutive equations have usually been formulated in terms of the Eulerian strain rate tensor $\mathbf{D}$ and an objective rate of stress (which measure of stress one should use will be discussed in the section about work conjugacy). The result can be split into a deviatoric tensor equation

$${}^{d}\dot{\boldsymbol{\sigma}}^{o} = 2\mu^{d}\mathbf{D} \tag{23}$$

(a superscripted prefix ^{d} denotes the deviatoric part of a tensor) and a scalar equation for the hydrostatic components

$$\dot{p} \equiv \frac{1}{3}\,\text{tr}\,(\dot{\boldsymbol{\sigma}}) = \frac{3\lambda + 2\mu}{3}\,\text{tr}\,(\mathbf{D}). \tag{24}$$

The scalar equation presents no difficulty because the time derivative of a scalar is always objective. However, with the classical choice of the objective stress rate as the J-rate or Z-rate, the tensor Eq. (23) does not represent an isotropic elastic material which, for the linear case, is defined by

$$\boldsymbol{\sigma} = \lambda\,\text{tr}\,(\mathbf{e})\mathbf{I} + 2\mu\mathbf{e}. \tag{25}$$

The appropriate strain measure $\mathbf{e}$ is logarithmic strain.

Metzger and Dubey (1987) suggested that the use of the modified hypoelastic constitutive relation

$${}^{d}\dot{\boldsymbol{\sigma}}^{o} = 2\mu^{d}\dot{\mathbf{e}}^{o} \tag{26}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

yields a solution which is identical with the isotropic elastic result

$$ {}^{d}\boldsymbol{\sigma} = 2\mu^{d}\mathbf{e}. \tag{27} $$

We can prove the identity by transforming (27) to the principal axes, on which follows:

$$ {}^{d}\boldsymbol{\Sigma} = 2\mu^{d}\mathbf{E} \Rightarrow {}^{d}\dot{\boldsymbol{\Sigma}} = 2\mu^{d}\dot{\mathbf{E}}. \tag{28} $$

Substituting into (26) yields

$$ {}^{d}\dot{\boldsymbol{\sigma}}^{o^{(e)}} - 2\mu^{d}\dot{\mathbf{e}}^{o^{(e)}} = ({}^{d}\dot{\boldsymbol{\Sigma}} - 2\mu^{d}\dot{\mathbf{E}}) - \boldsymbol{\Omega}_{\mathrm{rel}}^{o^{(e)}}({}^{d}\boldsymbol{\Sigma} - 2\mu^{d}\mathbf{E}) + ({}^{d}\boldsymbol{\Sigma} - 2\mu^{d}\mathbf{E})\boldsymbol{\Omega}_{\mathrm{rel}}^{o^{(e)}} = 0. \tag{29} $$

Hence, given the isotropic elastic relation (27), the hypoelastic formulation (26) holds. To prove the converse, we note that (26) is a system of quasi-linear differential equations. Furthermore, we know that (27) is a solution of this system and, by invoking the uniqueness theorem, we can conclude that the solution is unique if the given boundary conditions are compatible with (27).

Two important consequences follow from this result. First, the specific choice of the objective rate in (26) has no influence on the result. The isotropic-hypoelastic model is independent of this choice as long as the rates of stress and strain refer to the same frame of reference. Thus, the second consequence is that the classical formulations for which (29) does not hold cannot yield the correct result in general. From the work of Hoger (1986) and Gurtin and Spear (1983) follows that the classical form holds only when $\mathbf{R}^{E} = \mathbf{R}$, giving a good approximation of the solution when the deformation gradient differs from the unit matrix by only a small amount, and when its time derivative is small. Especially when the deformation involves large shearing, these conditions are too restrictive.

Finally, the result can be extended to nonlinear isotropic elastic materials described by

$$ \mathbf{e} = \mathbf{f}(I_1, I_2, I_3, \boldsymbol{\sigma}), \tag{30} $$

where $\mathbf{f}$ is an isotropic, tensor-valued function and I_1, I_2, and I_3 are the invariants of the stress tensor. The equivalent rate form is

$$ \dot{\mathbf{e}}^{o} = (\mathbf{f}(I_1, I_2, I_3, \boldsymbol{\sigma})) \cdot o. \tag{31} $$

For such a material, the proof (29) and its consequences remain valid.

For the plastic constitutive relation, we will follow the approach suggested by Dubey and Bedi (1989, 1991). We start from a uniaxial description of the stress-strain curve as

$$ e = f(\sigma) \tag{32} $$

for rate-independent plasticity, using the uniaxial (logarithmic) strain e and true stress σ. Similarly, a rate-sensitive material is described by the uniaxial relation

$$ \dot{e} = g(\sigma). \tag{33} $$

We introduce an equivalent stress and strain as the J_2-norm of the stress and the strain tensor:

$$ e_{eq} = \sqrt{\tfrac{2}{3}\,{}^{d}e_{ij}\,{}^{d}e_{ij}}, \tag{34} $$

and

$$ \sigma_{eq} = \sqrt{\tfrac{3}{2}\,{}^{d}\sigma_{ij}\,{}^{d}\sigma_{ij}}. \tag{35} $$

Note that these definitions let σ_{eq} and e_{eq} become identical with σ and e for a uniaxial deformation. The definition for $\dot{e}_{eq}$ follows from (34) as

$$ \dot{e}_{eq} = \frac{2}{3}\frac{{}^{d}\dot{e}_{ij}\,{}^{d}e_{ij}e}{e} = \frac{2}{3}\frac{{}^{d}\dot{e}_{ij}^{o}\,{}^{d}e_{ij}}{e}. \tag{36} $$

For the rest of this section, we will assume that the deviatoric parts of the stress and strain tensors have been taken and define

$$ e_{ij} \doteq {}^{d}e_{ij}, \quad \sigma_{ij} \doteq {}^{d}\sigma_{ij}. \tag{37} $$

We now generalize the uniaxial flow rules (32) and (33) to tensorial equations. Multiply the rate form by e to obtain

$$ \dot{e}e = \frac{f'(\sigma)\dot{\sigma}}{\sigma}\sigma e \quad \text{and} \quad \dot{e}e = \frac{g(\sigma)}{\sigma}\sigma e. $$

This is equivalent to

$$ \frac{2}{3}\dot{e}_{ij}^{o}e_{ij} = \frac{f'(\sigma_{eq})\dot{\sigma}_{eq}}{\sigma_{eq}}\frac{\sigma_{eq}e_{eq}}{\sigma_{kl}e_{kl}}\sigma_{ij}e_{ij} + \tilde{n}_{ij}e_{ij} $$

and

$$ \frac{2}{3}\dot{e}_{ij}^{o}e_{ij} = \frac{g(\sigma_{eq})}{\sigma_{eq}}\frac{\sigma_{eq}e_{eq}}{\sigma_{kl}e_{kl}}\sigma_{ij}e_{ij} + \tilde{n}_{ij}e_{ij} $$

where $\tilde{n}_{ij}e_{ij} = 0$. With the definitions

$$ H \equiv \frac{1}{f'(\sigma_{eq})}, \quad H^{*} \equiv \frac{1}{g(\sigma_{eq})}, \quad g \equiv \frac{\sigma_{ij}e_{ij}}{\sigma_{eq}e_{eq}} \tag{42} $$

the flow rules become

$$ \dot{e}_{ij}^{o} = \frac{3}{2}\frac{\dot{\sigma}_{eq}}{Hg\sigma_{eq}}\sigma_{ij} + n_{ij} \text{ and } \dot{e}_{ij}^{o} = \frac{3}{2}\frac{1}{H^{*}g\sigma_{eq}}\sigma_{ij} + n_{ij}. \tag{43} $$

To explore the properties of the tensor $\mathbf{n}$ we transform to the principal axes of strain, on which

$$ n_{ij}^{(e)}E_{ij} = 0. \tag{44} $$

The only nontrivial elements in this sum are those on the main diagonal ($i = j$). Since $\mathbf{e}$ is a deviatoric tensor, n_{ij} could be equal to $k\delta_{ij}$ (a multiple of the unit tensor) but this is ruled out by the deviatoric character of the whole equation. Another choice which can be made in general is to let the diagonal elements of $\mathbf{n}$ vanish on the principal axes of $\mathbf{e}$:

$$ n_{ii}^{(e)} = 0 \quad \text{(no sum)}. \tag{45} $$

At the same time, $\mathbf{n}$ must be symmetric and objective to maintain the correct mathematical and physical structure of the flow rule. It is easily demonstrated that the spin term $(\mathbf{P}^{o})^{1/2}(\mathbf{D} - \dot{\mathbf{e}}^{E})(\mathbf{P}^{o})^{1/2}$ of the generalized objective rate (18) is always an appropriate choice for $\mathbf{n}$. Thus, the tensor $\mathbf{n}$ allows one to express a flow rule which has been set up with a specific objective rate in terms of any other objective rate. This result is significant because it establishes a property similar to the frame-invariance of the hypoelastic model which was shown earlier.

If kinematic hardening is included, a procedure similar to that used for the flow rule can be applied to the evolution equation for the tensorial back stress α. We start from the uniaxial relation

$$ \dot{\alpha} = C_1\dot{e} - C_2|\dot{e}|\alpha \tag{46} $$

where C_1 and C_2 are material constants. We introduce the J_2-norm for α as

$$ \alpha_{eq} = \sqrt{\frac{3}{2}\alpha_{ij}\alpha_{ij}}\,; \quad \dot{\alpha} = \frac{3}{2}\frac{\dot{\alpha}_{ij}^{o}\alpha_{ij}}{\alpha_{eq}}. \tag{47} $$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The generalization process leads then to the following system which describes the material behavior:

$$\dot{e}^o_{ij} = \frac{3}{2}\frac{\dot{\sigma}_\alpha}{gH\sigma_\alpha}(\sigma_{ij} - \alpha_{ij}) + n_{ij}$$

$$\dot{\alpha}^o_{ij} = \frac{2}{3}\frac{C_1}{h}\dot{e}^o_{ij} - C_2|\dot{e}|\alpha_{ij} + m_{ij};$$

$$\sigma_\alpha = \sqrt{\frac{3}{2}(\sigma_{ij} - \alpha_{ij})(\sigma_{ij} - \alpha_{ij})},$$

$$g = \frac{(\sigma_{ij} - \alpha_{ij})e_{ij}}{\sigma_\alpha e_{eq}},$$

$$h = \frac{3}{2}\frac{e_{eq}}{\alpha_{eq}}\frac{\dot{e}^o_{ij}\alpha_{ij}}{\dot{e}^o_{ij}e_{ij}}. \tag{48}$$

Note that the first equation is the flow rule, modified for kinematic hardening. For simplicity, only the rate-independent form was given. In practice, the system (48) is often supplemented by an evolution equation for the isotropic hardening parameter H. The factors g and h are generated by the generalization from the uniaxial (scalar) to the tensorial form of the constitutive equations. If the rate of strain is not coaxial with the back stress α, the evolution equation is corrected by the factor h. The tensor $\mathbf{m}$ must satisfy the condition $m_{ij}\alpha_{ij} = 0$. Therefore, any objective rate of the back stress used in the evolution equation can be transformed to any other objective rate without changing the result. The objective rate of strain used in the evolution equation should be the same as that used in the flow rule. A logical (though not mandatory) choice for the rate of the back stress is to employ the same rate as for the strain.

Work Conjugacy

The rate of change of internal elastic energy stored in a body equals the volume integral

$$\dot{W} = \int_V \tau_{ij}D_{ij}d\mathbf{x} = \int_V \tau_{ij}\dot{e}^D_{ij}d\mathbf{x}. \tag{49}$$

Equation (49) establishes work conjugacy in the sense of Hill (1978) between the Kirchhoff stress $\boldsymbol{\tau} = \det(\mathbf{V})\cdot\mathbf{s}$, where $\mathbf{s}$ is Cauchy or true stress, and the logarithmic strain measure $\mathbf{e}$. Note that in the general case, this conjugacy exists in (49) only for a special rate of strain, here the D-rate. However, the same is true for other work-conjugate pairs, like the second Piola-Kirchhoff stress and Green's strain tensor.

For an isotropic-elastic material described by the relations (26) and (27), the elastic energy becomes

$$\dot{W} = \int_V \tau_{ij}\dot{e}^D_{ij}d\mathbf{x} = \int_V (2\mu e_{ij}\dot{e}^D_{ij} + 3\lambda e_{ii}\dot{e}^D_{jj})d\mathbf{x}$$

$$= \int_V (2\mu\sum_i E_i D^{(e)}_{ii} + 3\lambda\sum_i E_i\sum_i D^{(e)}_{ii})d\mathbf{x}. \tag{50}$$

In fact, since only the diagonal elements of $\mathbf{D}$ enter this equation which are the same for all objective rates of strain of the type (18), any of these rates could be used instead of $\mathbf{D}$ in (50).

In a rigid-plastic body, the deformation energy is dissipated completely. The dissipation function for such a process is

$$\dot{\Psi} = \int_V \sigma_{ij}D_{ij}d\mathbf{x} \quad \text{or} \quad \dot{\Psi} = \int_V (\sigma_{ij} - \alpha_{ij})D_{ij}d\mathbf{x}, \tag{51}$$

where the second form is applicable if kinematic hardening is present. Using the flow rule

$$\dot{e}^o_{ij} = \dot{\phi}\sigma_{ij} \quad \text{or} \quad \dot{e}^o_{ij} = \dot{\phi}(\sigma_{ij} - \alpha_{ij}), \tag{52}$$

the dissipation function for both models can be written in terms of the strain rate as

$$\dot{\Psi} = \int_V \dot{\phi}\dot{e}^o_{ij}D_{ij}d\mathbf{x} = \int_V \dot{\phi}\dot{e}^o_{ij}\dot{e}^D_{ij}d\mathbf{x}. \tag{53}$$

Now we can employ the definition of the generalized objective rate (10) to obtain

$$\dot{e}^o_{ij}D_{ij} = \dot{e}^{o(e)}_{ij}D^{(e)}_{ij}$$

$$= \sum_i \dot{D}^{(e)2}_{ii} + 2(p^o_{12}D^{(e)2}_{12} + p^o_{13}D^{(e)2}_{13} + p^o_{23}D^{(e)2}_{23}). \tag{54}$$

Since we assume that a deformation is taking place, the sum of the squared diagonal terms must be positive. The sign of the second sum in (54) depends on the signs of the factors p^o. The total dissipation is required to be positive, and for some deformations the off-diagonal elements of $\mathbf{D}^{(e)}$ can become large compared to the diagonal elements. Hence, the condition

$$p^o_{ij} \geq 0 \tag{55}$$

must be imposed on the p^o for any objective strain rate which is to be used in the flow rule (52). Thus, the E-rate is found to be the limiting case of admissible strain rates, since it implies $p^E = 0$. Due to volume constancy, the distinction between Cauchy stress and Kirchhoff stress vanishes in a rigid-plastic material, and $\boldsymbol{\sigma}$ can be identified with either.

Application

To demonstrate the outlined approach, we show its application to the simple shear deformation. Consider a plate with constant thickness which is sheared such that its lateral dimensions remain the same, Fig. 1. The deformation gradient for this deformation is $\mathbf{F} = \mathbf{g}_1 \otimes \mathbf{g}_1 + \mathbf{g}_2 \otimes \mathbf{g}_2 + \gamma\mathbf{g}_1 \otimes \mathbf{g}_2$, where the vectors $\mathbf{g}_i$ form the fixed basis. From this, the principal values of logarithmic strain are obtained as

$$E_1 = -E_2 = \operatorname{arsinh}\left(\frac{\gamma}{2}\right), \tag{56}$$

and the orientation of the principal axes is found to be

$$\beta = \frac{\pi}{4} - \frac{1}{2}\arctan\frac{\gamma}{2}. \tag{57}$$

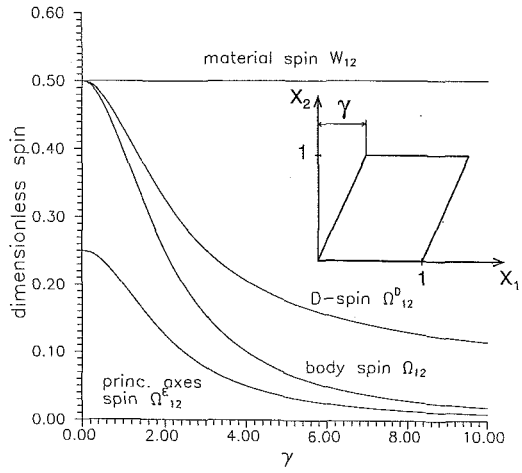


Fig. 1 Dimensionless deformation-related spins $\Omega^o_{12}/\dot{\gamma}$ as a function of the shear displacement γ in simple shear

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The deformation is purely deviatoric. The components of the Eulerian strain rate $\mathbf{D}$ on the fixed axes and on the principal axes of strain are

$$D_{11} = D_{22} = 0, \quad D_{12} = \frac{\dot\gamma}{2};$$

$$D_{11}^{(e)} = -D_{22}^{(e)} = \frac{\dot\gamma}{2}\sin 2\beta, \quad D_{12}^{(e)} = \frac{\dot\gamma}{2}\cos 2\beta. \tag{58}$$

The nonzero fixed-axes components of body spin $\mathbf{\Omega}$, material spin $\mathbf{W}$, D-spin $\mathbf{\Omega}^D$ and principal axes spin $\mathbf{\Omega}^E$ are

$$\Omega_{12} = \frac{2\dot\gamma}{4+\gamma^2}, \quad W_{12} = \frac{\dot\gamma}{2}, \quad \Omega_{12}^E = \frac{\dot\gamma}{4+\gamma^2},$$

$$\Omega_{12}^D = \frac{\gamma\sqrt{4+\gamma^2} + 4\,\text{arsinh}\left(\frac{\gamma}{2}\right)}{4(4+\gamma^2)\,\text{arsinh}\left(\frac{\gamma}{2}\right)}\dot\gamma. \tag{59}$$

By integration we obtain the corresponding rotations β as

$$\beta_Z = -\arctan\frac{\gamma}{2}, \quad \beta_J = -\frac{\gamma}{2},$$

$$\beta_D = -\frac{1}{2}\arctan\frac{\gamma}{2} - \frac{1}{2}\,\text{Sih}\left(\text{arsinh}\left(\frac{\gamma}{2}\right)\right), \tag{60}$$

and $\beta_E = \beta$ from (57). The function Sih (x) is defined as

$$\text{Sih}\,(x) \equiv \int_0^x \frac{\sinh\xi}{\xi}\,d\xi = \sum_{i=0}^{\infty}\frac{x^{2i+1}}{(2i+1)\cdot(2i+1)!}. \tag{61}$$

Figures 1 and 2 show the spins and rotations as a function of the shear displacement γ. The graphs reveal clearly the ordering established by (20). Note that β_E and β_Z are bounded; β_D is very slowly increasing for large γ, and β_J increases linearly with γ. Much of the observed behaviour of constitutive models using these rates can already be inferred from this behaviour of the rotations (and spins).

To compare the hypoelastic solution from the model proposed here with the solutions from other models (see also the discussion by Dienes, 1979), consider an equation of the form

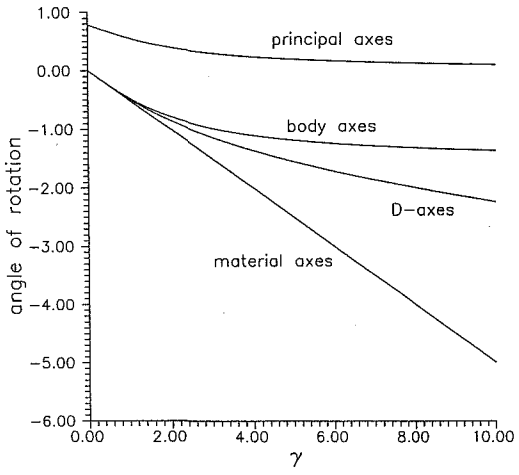


Fig. 2 Rotation angles of various spinning frames in simple shear (see Eqs. (60))

$$D_{ij} = \frac{1}{2\mu}\,\mathring{\sigma}_{ij}^{o}. \tag{62}$$

Let $\omega_0 = \Omega_{12}^o$ be the nonzero component of the spin of the objective rate. The system of differential equations that results from (62) is then

$$\sigma_{11}' - 2\frac{\omega_0}{\dot\gamma}\sigma_{12} = 0$$

$$\sigma_{12}' + 2\frac{\omega_0}{\dot\gamma}\sigma_{11} = \mu \tag{63}$$

(a prime denotes differentiation w.r.t. γ), where the deviatoric character of stress has been exploited. A general solution of this system is found to be

$$\begin{pmatrix}\sigma_{11}\\ \sigma_{12}\end{pmatrix} = \begin{pmatrix}C_1\cos 2\beta_0 + C_2\sin 2\beta_0\\ C_1\sin 2\beta_0 - C_2\cos 2\beta_0\end{pmatrix}$$

$$+\mu\begin{pmatrix}\left(\int_0^\gamma \sin 2\beta_0 d\gamma\right)\cos 2\beta_0 - \left(\int_0^\gamma \cos 2\beta_0 d\gamma\right)\sin 2\beta_0\\ \left(\int_0^\gamma \cos 2\beta_0 d\gamma\right)\cos 2\beta_0 + \left(\int_0^\gamma \sin 2\beta_0 d\gamma\right)\sin 2\beta_0\end{pmatrix}$$

$$\beta_0 = \int_0^\gamma\left(-\frac{\omega_0}{\dot\gamma}\right)d\gamma. \tag{64}$$

The rotations β_0 for J-rate, Z-rate, and D-rate are given by (60). Note that the terms in the first parenthesis on the r.h.s. of (64) represent the homogenous part of the solution. This part oscillates if β_0 exceeds $\pi/2$. Therefore, oscillations are possible for the J-rate and D-rate of stress for which β_0 is unbounded, but not for the Z-rate. However, the total solution also includes the inhomogeneous part (second set of terms on the r.h.s. of (64)) which may be nonoscillatory. Thus, depending on the initial conditions, the solution may be monotonic even if β_0 is unbounded. For an elastic material we require $\sigma_{11} = \sigma_{12} = 0$ as $\gamma = 0$.

The classical J-rate model leads to the solution

$$\sigma_{11} = \mu(1-\cos\gamma)$$

$$\sigma_{12} = \mu\sin\gamma. \tag{65}$$

Here the oscillation of the homogenous part carries over into the total solution.

The Z-rate model which has been suggested to eliminate the oscillation (Dienes (1979)) gives

$$\sigma_{11} = 2\mu\,\frac{\gamma\left(2\arctan\frac{\gamma}{2} - \frac{\gamma}{2}\right) - \left(1-\frac{\gamma^2}{4}\right)\ln\left(1+\frac{\gamma^2}{4}\right)}{1+\frac{\gamma^2}{4}}$$

$$\sigma_{12} = 2\mu\,\frac{\gamma\ln\left(1+\frac{\gamma^2}{4}\right) + \left(1-\frac{\gamma^2}{4}\right)\left(2\arctan\frac{\gamma}{2} - \frac{\gamma}{2}\right)}{1+\frac{\gamma^2}{4}}. \tag{66}$$

The model is successful in avoiding oscillations. However, the solution does not agree with that of the isotropic elastic model (25).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Use of the D-rate is consistent with the approach proposed here and should yield the correct isotropic-elastic solution. Indeed, we find

$$\sigma_{11} = 2\mu \frac{\gamma}{\sqrt{4+\gamma^2}} \operatorname{arsinh}\left(\frac{\gamma}{2}\right)$$

$$\sigma_{12} = 4\mu \frac{1}{\sqrt{4+\gamma^2}} \operatorname{arsinh}\left(\frac{\gamma}{2}\right), \tag{67}$$

which also follows from (25), (56), and (57). The initial conditions eliminate the oscillatory homogenous part of the general solution. A graphical comparison of the three solutions which were derived here can be found in the paper by Metzger and Dubey (1987).

Next, consider a rigid-plastic deformation with isotropic hardening. We assume the stress-strain curve is of the Ramberg-Osgood type

$$e_{eq} = K\sigma_{eq}^n, \quad e_{eq} = \frac{2}{\sqrt{3}} \operatorname{arsinh}\left(\frac{\gamma}{2}\right). \tag{68}$$

The flow rule is given by (43) with $\mathbf{n} \equiv 0$ which was shown to be compatible with the uniaxial representation. We obtain the system

$$\dot{e}_{11}^{o(e)} = D_{11}^{(e)} = \frac{3}{2}\frac{\dot{\sigma}_{eq}}{gH\sigma_{eq}}\sigma_{11}^{(e)}$$

$$e_{12}^{o(e)} = p^o D_{12}^{(e)} = \frac{3}{2}\frac{\dot{\sigma}_{eq}}{gH\sigma_{eq}}\sigma_{12}^{(e)}. \tag{69}$$

$\mathbf{D}^{(e)}$ has been determined in (58), and the p^o are given by (11)–(14). The parameters g and H are defined in (42). To derive the solution of (69) we divide the first equation by the second, and substitute into the definition for σ_{eq}, giving

$$\sigma_{eq} = \sqrt{3}\sigma_{12}^{(e)}\sqrt{\left(\frac{\sigma_{11}^{(e)}}{\sigma_{12}^{(e)}}\right)^2 + 1} = \sqrt{3}\sigma_{12}^{(e)}\sqrt{\left(\frac{\tan 2\beta}{p^o}\right)^2 + 1}. \tag{70}$$

The last two equations can be solved for the components of $\boldsymbol{\sigma}$ on the principal axes of strain. Transforming back to the fixed axes, the result is

$$\sigma_{11} = \frac{\sigma_{eq}}{\sqrt{3}}\frac{(1-p^o)\sin 2\beta \cos 2\beta}{\sqrt{\sin^2 2\beta + p^{o^2}\cos^2 2\beta}}$$

$$\sigma_{12} = \frac{\sigma_{eq}}{\sqrt{3}}\frac{\sin^2 2\beta + p^o \cos^2 2\beta}{\sqrt{\sin^2 2\beta + p^{o^2}\cos^2 2\beta}},$$

$$\sigma_{eq} = \left(\frac{\frac{2}{\sqrt{3}}\operatorname{arsinh}\left(\frac{\gamma}{2}\right)}{K}\right)^{1/n}, \tag{71}$$

where (68) has been utilized to eliminate σ_{eq}. The trigonometric expressions are defined as

$$\sin 2\beta = \frac{2}{\sqrt{4+\gamma^2}}, \quad \cos 2\beta = \frac{\gamma}{\sqrt{4+\gamma^2}}. \tag{72}$$

Several important properties of the solution are determined by the character of p^o. First, we note that $p = 1$ implies that the normal stress is identically zero. Hence, an isotropic flow rule using the D-rate predicts no higher-order effects in shear, a well-known fact. Whatever the value of p is, because $\beta = 45$ deg when $\gamma = 0$ and because σ_{eq} follows the uniaxial stress-strain curve with $e_{eq} \approx \gamma/\sqrt{3}$ for small γ, all admissible rates reduce correctly to the theory of infinitesimal deformations. The normal stresses are at least of second order in the shear displacement when this displacement is small. In fact, for the Z-rate and the J-rate which start with $p = 1$, the predicted normal stresses are of even higher order. The normal stresses, therefore, develop only slowly for these two rates. Observe also that if $p < 1$, a tensile stress is predicted in the 1-direction and, because the stress tensor is deviatoric, a compressive stress occurs in the 2-direction. For $p > 1$, the sign of these stresses is interchanged. If these results are translated into the behavior in a torsion test, $p < 1$ means that the specimen has a tendency to increase its length in the direction of the axis of twist, paired with a reduction of the cross section. The Z-rate model and the E-rate model show this behavior. Conversely, if $p > 1$ as in the J-rate model, torsion leads to axial shortening and radial expansion. The stress predictions for simple shear are shown in Fig. 3.

It was demonstrated how a flow rule with isotropic hardening yields no higher-order effects in shear if the D-rate is employed. However, with kinematic hardening these effects can be introduced. In keeping with our previous derivations, we will consider the system

$$\dot{e}_{ij}^D = D_{ij} = \frac{3}{2}\frac{\dot{\alpha}_\alpha}{gH\sigma_\alpha}(\sigma_{ij} - \alpha_{ij}),$$

$$\dot{\alpha}_{ij}^D = \frac{2}{3}\frac{C_1}{h}D_{ij} - C_2 e_{eq}\alpha_{ij}. \tag{73}$$

To keep the example simple and illustrative, we assume that no isotropic hardening occurs, thus $\sigma_\alpha = \sigma_y$, the yield stress of the material. Using (48), we can derive for the stress

$$\sigma_{11} = -\sigma_{22} = \alpha_{11}, \quad \sigma_{12} = \frac{\sigma_y}{\sqrt{3}} + \alpha_{12}. \tag{74}$$

With the definitions of the strain tensor, the factor h is determined as

$$h = \frac{\alpha_{12}}{\sqrt{\alpha_{11}^2 + \alpha_{12}^2}\sin 2\beta}. \tag{75}$$

The evolution equations for the back stress components become

$$\alpha_{11}' + \beta_D\alpha_{12} = -\frac{2}{\sqrt{3}}\frac{C\alpha_{11}}{\sqrt{4+\gamma^2}}$$

$$\alpha_{12}' - \beta_D\alpha_{11} = \frac{2}{3}\frac{CK_\alpha\alpha_{12}}{\sqrt{\alpha_{11}^2+\alpha_{12}^2}\sqrt{4+\gamma^2}} - \frac{2}{\sqrt{3}}\frac{C\alpha_{12}}{\sqrt{4+\gamma^2}}. \tag{76}$$

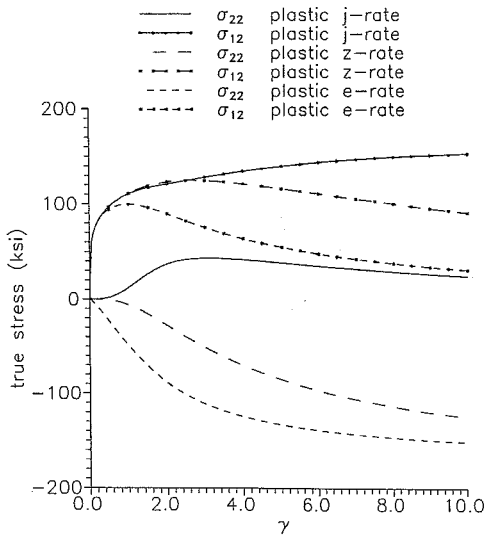


Fig. 3 Normal and shear stresses in simple shear predicted by the rigid-plastic model with isotropic hardening. Results are obtained using different objective rates in (43). The underlying uniaxial model is Ramberg-Osgood with K = (218 ksi)$^{-4.804}$ and n = 4.808 in (68).

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

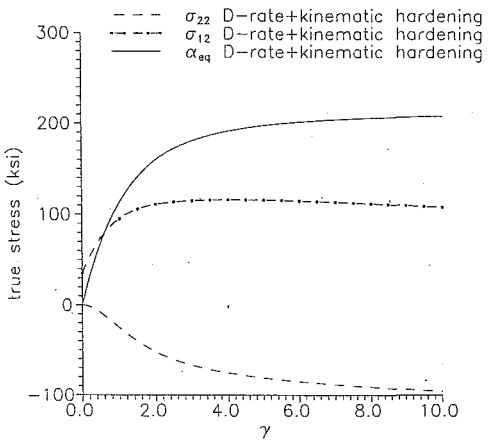


Fig. 4 Normal and shear stress in simple shear resulting from the kinematic hardening model (48). Material parameters used were σ_y = 61.3 ksi, C = 1.35 and K_α = 214 ksi in (76). Note that the α_{eq}-γ curve calculated from the tensor components of α is identical with the uniaxial curve (46).

The rotation β_D is given in (60). The equations are nonlinear, and we are not aware of an analytical solution. Therefore, the system was solved numerically. Figure 4 shows a plot of the stresses as a function of the shear displacement γ. The parameter values were σ_y = 61.3 ksi, K_α = 214 ksi and C = 1.35, identical with the set used by Zbib and Aifantis (1987). The behavior of the stress components resembles that observed for the Z-rate. Since most experimental results support the development of compressive axial stresses in torsion, the curves seem reasonable. Due to the derivation of the evolution equations, the α_{eq}-γ curve is identical with that resulting from the uniaxial relation (46) with $C_1 = K_\alpha C$ and $C_2 = C$.

Conclusions

The unified form of deformation-related spin tensors which is presented in this article gives rise to a more general concept of a spin. Some mathematical and thermodynamical restrictions on this generalized spin have been discussed. The properties and interrelations of spins which have been traditionally used in kinematical modelling emerge very clearly from the formalism. The existence of a D-spinning frame could be established on which the rate of logarithmic strain equals the Eulerian strain rate. The formalization also allows constitutive relations to be derived for a whole class of objective rates. Properties of the solution can be inferred from the structure of the spin parameter without actually solving the equations. Conversely, if the desired behavior of the solution is known, one can draw conclusions about the necessary character of the spin parameter.

The objective D-rate permits a classical hypoelastic constitutive formulation for a material which is at the same time isotropic elastic. It was also shown how a slightly modified hypoelastic relation yields this result for rates based on any corotational frame. Shortcomings in traditional rate formulations, like oscillation in the simple shear solution, are, therefore, attributable to different frames of reference used for the rates of stress and strain in these relations.

A slightly modified plastic flow rule which employs an arbitrary rate of strain, rather than only the D-rate, predicts higher-order effects in shear without need for a back stress tensor. The expected behavior of normal stresses or strains in a torsion test (which is a higher-order effect) is shown to be related to the departure of the spin parameter from unity. However, a comparison with experiments reveals that for the traditional rates of strain, like the Z-rate or J-rate, these effects develop too slowly (fourth-order) at small strains, and become too large at higher strains. Furthermore, because these rates are based on kinematics, they cannot account for the distinct variations in the response of different materials. These observations suggest that an appropriate rate should contain a material-dependent parameter. With the help of the general form of spins and corotational rates proposed in this article, the selection of an acceptable rate is greatly simplified.

When kinematic hardening is introduced into the flow rule, the effect is similar to what is observed from the modified flow rule with isotropic hardening. The approach outlined in this article ensures that both the flow rule and the evolution equation of the back stress reduce to the uniaxial relations in a consistent way, both tensorially and in terms of equivalent stress and strain. What has been said about the objective rate in the previous paragraph applies here as well, but the introduction of the back stress tensor $\boldsymbol{\alpha}$ allows the construction of other objective rates based on this tensor. Such rates have been discussed in the literature, e.g., by Lee et al. (1983) and Bammann and Aifantis (1987).

Acknowledgments

This work was supported by the Natural Science and Engineering Research Council of Canada. The authors are indebted to M. Singh for a careful revision of the manuscript.

References

Bammann, D. J., and Aifantis, E. C. A., 1987, "A Model for Finite-Deformation Plasticity," *Acta Mechanica,* Vol. 69, pp. 97–117.

Dafalias, Y. F., 1983, "Corotational Rates for Kinematic Hardening at Large Plastic Deformation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 561–565.

Dienes, J. K., 1979, "On the Analysis of Rotation and Stress Rate in Deforming Bodies," *Acta Mechanica,* Vol. 32, pp. 217–232.

Dubey, R. N., 1987, "Choice of Tensor Rates: A Methodology," *Solid Mechanics Archives,* Vol. 24, No. 4, pp. 233–244.

Dubey, R. N., and Bedi, S., 1989, "Some Implications of the MDB Flow Rule in Plasticity," *Advances in Plasticity,* Proceedings of PLASTICITY '89, A. S. Khan and N. Tokuda, eds., Pergamon Press, New York, pp. 291–295.

Dubey, R. N., and Bedi, S., 1991, "On Flow Rules in Plastic Deformation," *Archives of Mechanics,* Vol. 43, No. 1, pp. 129–137.

Green, A. E., and McInnis, B. C., 1967, "Generalized Hypo-Elasticity," *Proceedings of the Royal Society Edinburgh,* Vol. A67, pp. 220–230.

Halleux, J. P., and Donea, J., 1986, "A Discussion of Cauchy Stress Formulations for Large Strain Analysis," *Finite Element Method for Nonlinear Problems,* Proc. Europe-US Symp. Trondheim, Norway., P. G. Bergan, K. J. Bathe, and W. Wunderlich, eds., Springer, Berlin.

Hill, R., 1978, "Basic Aspects of Invariance in Solid Mechanics," *Advances in Applied Mechanics,* Vol. 18, C. S. Yih, ed., Academic Press, New York, pp. 1–75.

Lee, E. H., Mallet, R. L., and Wertheimer, T. B., 1983, "Stress Analysis for Anisotropic Hardening in Finite Deformation Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 554–560.

Mehrabadi, M. M., and Nemat-Nasser, S., 1987, "Some Basic Kinematical Relations for Finite Deformations of Continua," *Mechanics of Materials,* Vol. 6, pp. 127–138.

Metzger, D. R., and Dubey, R. N., 1987, "Corotational Rates in Constitutive Modelling of Elastic-Plastic Deformation," *International Journal of Plasticity,* Vol. 4, pp. 341–368.

Nagtegaal, J. C., and De Jong, J. E., 1982, "Some Aspects of Nonisotropic Work Hardening in Finite-Strain Plasticity," *Proceedings of the Workshop on Plasticity of Metals at Finite Strain: Theory, Experiment and Computation,* E. H. Lee, and R. L. Mallet, eds., R. P. I. Troy, pp. 65–102.

Ogden, R. W., 1984, *Non-linear Elastic Deformations,* (Ellis Horwood Series in Mathematics and its Applications), Halsted Press, New York.

Truesdell, C., and Noll, W., 1965, *The Non-linear Field Theories in Mechanics,* Handbuch der Physik, Bd. III/3, S. Flügge, ed., Springer-Verlag, Berlin.

Zbib, H. M., and Aifantis, E. C., 1987, "The Concept of Relative Spin and its Implications to Large Deformation Theories," *Mechanics of Microstructures,* MM Report No. 13, Michigan Technological University, Houghton, MI.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

R. Becker
Assoc. Mem. ASME.

M. E. Karabin
Mem. ASME.

J. C. Liu
Mem. ASME.

Alcoa Technical Center,
100 Technical Drive,
Alcoa Center, PA 15069

R. E. Smelser
University of Idaho,
Mechanical Engineering Department,
Moscow, ID 83844-0902
Mem. ASME.

Distortion and Residual Stress in Quenched Aluminum Bars

Predictions of distortion and residual stress for rectangular bars, quenched on one surface, are compared with experimental measurements. The heat flow is unidirectional, and the dominant stress component acts along the length of the bars. Thermocouples measure the transient temperature distribution, and displacement transducers monitor the distortion of the bars into circular arcs. The thermal history inferred from an inverse heat transfer calculation drives the mechanical stress and deformation analysis. The numerical analyses demonstrate the sensitivity of the quench distortion to the thermal boundary conditions, the thermophysical properties and the constitutive model. Agreement between the predicted and measured distortion is only qualitative. The predicted residual stress field compares very well with the residual stress inferred from layer removal techniques.

1 Introduction

Many structural materials are heat treated in order to achieve desired properties. Frequently this involves holding the material at an elevated temperature and then quenching it in oil, water, or some other medium. The temperature gradients caused by the quench can be severe, and the thermal strain gradients can induce inelastic strains. Heat treated parts often distort during quenching and residual stresses are introduced. The residual stresses can lead to distortion during subsequent machining. They can also be deleterious by affecting fracture and corrosion performance. Problems with quench distortion, machine distortion, and residual stress are common, affecting castings, forgings, extrusions, and plate products.

Models predicting residual stresses developed during heat treatment have been available for many years (Landau et al., 1958). More recently, predictions of residual stress in quenched plate have been verified by comparison with experiments (Fletcher and Price, 1981; Fletcher and Soomro, 1986; Jeanmart and Bouvaist, 1984). Many of the predictions are in good agreement with the measurements, and some of the results (Boyer and Boivin, 1985) seem to suggest that the residual stress is relatively insensitive to moderate variations in material properties and process parameters. However, a study by Fletcher and Price (1981) illustrates that the agreement between model and experiment can vary greatly.

The present work seeks to gain a better understanding of quench distortion and residual stress development. An experiment proposed by Aksel (1990) is used. The experiment consists of a rectangular bar quenched on one surface. The heat flow is unidirectional and the bar distorts into a circular arc. The quench distortion of the bars is very sensitive to the material model and thermal history. This makes the test a good choice for model validation. A simulation of the experiment in the form of a beam model was performed by Aksel (1990) and Aksel et al. (1992). They compared the predicted final curvature of the bars to the measured curvature and found only qualitative agreement. The present work monitors the evolution of the distortion during the quench (Becker et al., 1994). The finite element method is more flexible than the previous analysis, but the results are substantially the same (Arthur et al., 1994).

2 Quench Experiment

Rectangular bars of commercial purity aluminum and Al-4.6 Mg approximately 25 mm by 75 mm by 320 mm (Fig. 1) were tested. For convenience the 320 mm direction (X_1) will be denoted the longitudinal direction or length; the 75 mm direction (X_2) will be called the height and the 25 mm direction the width. The bars were held in a furnace at 560°C for 20 minutes and quenched by immersing one 25 mm by 320 mm surface approximately 3 mm deep in a reservoir of water. Insulation was placed on the 75 mm by 320 mm sides of the bars to minimize heat transfer to the air. The heat transfer in the bars was essentially unidirectional, and the bars warped into circular arcs.

The specimens were set in a fixture (Fig. 1), which also holds a set of displacement transducers to monitor the distortion. The top of the frame is removable to facilitate inserting the hot bars. The specimens rest on a pair of knife edge supports 290 mm apart. The fixture sits in a reservoir of water with the surface of the water being about 1 mm below the top of the supports prior to the quench. The bars were removed from the furnace and placed in the quench rig. The top of the frame was replaced, and the water level in the tank was raised approximately 4 mm. Data were collected at 0.05-sec intervals.

Eight Linear Variable Differential Transformers (LVDT) in the top of the fixture measure the deflection. Two sensors are directly over the supports and the others are placed at 36.25 mm intervals toward the center. There is no transducer at the center. The transducers over the supports measure only the thermal contraction of the specimen.

The temperature distribution in the bars was monitored by five thermocouples. The locations are shown schematically in Fig. 1. The thermocouples entered the specimens from the 25 mm by 320 mm face opposite the quenched surface. Three thermocouples were inserted to a depth of 6.35 mm from the quenched surface at various locations along the length of the bar. The remaining two were placed at 25.4 mm and 63.5 mm from the quenched surface. The thermocouple junctions were

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 16, 1995; final revision, Mar. 4, 1996. Associate Technical Editor: W. N. Sharpe, Jr.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

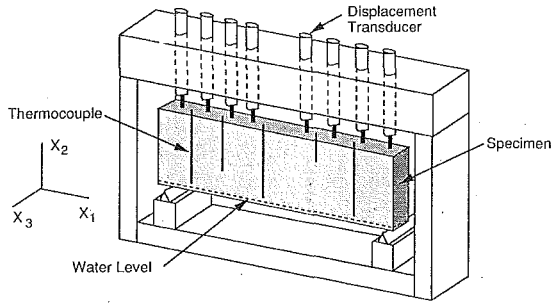


Fig. 1 Quench fixture and specimen

formed by contact of the two stripped leads with the bottom of the hole to avoid problems with thermal contact resistance and thermal mass.

2.1 Specimen Preparation. Tests were conducted on two alloys. One was commercial purity aluminum with 0.08 percent Si and 0.15 percent Fe, and the other was the same base material with an additional 4.6 percent Mg (Al-4.6 Mg). These materials were chosen because their strengths are significantly different and the properties will not vary significantly with time at elevated temperature. Consequently, the flow strength in the domain of interest can be modeled as a function of the temperature, strain rate, and a single state variable.

The specimens were held at temperature (560°C) for approximately 20 minutes prior to the quench. Both alloys were completely recrystallized before quenching. The commercial purity bars had a grain size of approximately 150 μm throughout, and the Al-4.6 Mg alloy had a 50 μm grain size at the center and grains in excess of 1000 μm near the surfaces.

2.2 Experimental Results.

2.2.1 Thermal. The thermocouple readings were slightly different (~5°C) at the elevated temperature even after the prolonged time in the furnace. The temperatures were adjusted linearly with temperature from zero adjustment at the room temperature calibration point to the uniform initial temperature 560°C.

The temperature histories for the commercial purity alloy are shown in Fig. 2. The histories for the Al-4.6 Mg are similar. There was some heat loss between the time the specimens were removed from the furnace and the beginning of the quench. The similarity of the temperatures at 6.35 mm from the surface indicates that the quench was uniform across the bars. There was, however, a small persistent difference for one of the thermocouples in the Al-4.6 Mg alloy. This may have been related to variations in the film boiling on the surface or perhaps to inaccurate placement of the thermocouple.

Comparison of the temperature histories at 63.5 mm from the quenched surface for the two alloys showed that the commercial purity material cooled faster. This resulted from the lower thermal conductivity of Al-4.6 Mg alloy which had alloying elements in solid solution.

2.2.2 Distortion. In the idealized situation of unidirectional heat transfer in a long, thin bar, the specimen would distort into a circular arc where the distortion could be characterized by a single parameter: the curvature. The curvature of the bars was calculated assuming that the bars form circular arcs. The distortion was also assumed to be symmetric about the midpoint, and the curvature is determined adequately from

$$\frac{1}{R} = 2\left(\frac{\delta_B - \delta_A}{d_B^2 - d_A^2}\right). \tag{1}$$

Here δ_A denotes the displacement measured by the transducer located at a distance d_A from the center of the bar. Curvature is positive when the center of curvature is above the specimen. The displacement data from symmetric locations are averaged for these calculations.

Since data are available from four locations, six curvature histories can be constructed. For the commercial purity material, the average of these curvatures at the end of the test was 0.07 m^{-1} and the range was 0.01 m^{-1}. There was no consistent trend to indicate that the deviation was due to end effects. The curvature was also calculated from a least squares fit of a circular arc to the data. The results from this regression fit and that of the average of the curvatures determined from Eq. (1) were nearly indistinguishable. The averages of the six curvature calculations are shown in Fig. 3. The results from the duplicate tests are included to illustrate the magnitude of the test-to-test variation. The variation between duplicate tests and between the two alloys shows the extreme sensitivity of this experiment to changes in material properties and to natural variations in boundary conditions.

2.2.3 Residual Stress. A residual curvature indicates that the specimens deformed inelastically during quenching and that the bars contain residual stresses. The dominant residual stress component acts along the longitudinal (X_1) direction of the bars. The bars are long enough that these stresses do not vary along the length in the center portion of the specimen. Equilibrium requires that the shear stress and the stress in the X_2 direction be negligible in this central region. The 25 mm width is assumed to be too small to develop appreciable stress in the X_3 direction.

The residual stress was determined for the Al-4.6 Mg alloy only. The commercial purity material creeps at room temperature. During several months at room temperature, these specimens lost some of their curvature, and the residual stress would have relaxed as well.

Residual stresses in the Al-4.6 Mg specimens were inferred from a standard layer removal technique. A strain gage was

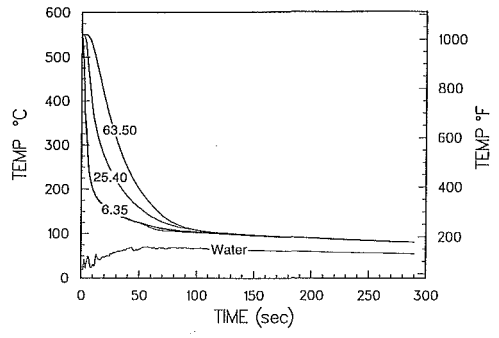


Fig. 2 Temperature history at three thermocouple locations for the commercial purity aluminum bar

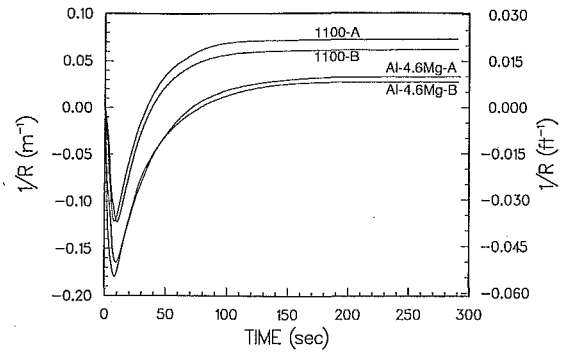


Fig. 3 Curvature histories for duplicate specimens of the two alloys

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

mounted on the center of the unquenched 25 mm by 320 mm surface. 1.27 mm thick layers of material were removed from the quenched surface. Removal of the stressed material caused the remainder of the specimen to distort. The elastic properties, strain gage readings and the thickness of the machined bar were used to calculate the stress in the layer that was removed. It was assumed that removal of a layer of material does not induce plastic deformation in the specimen. The results will be presented with the model predictions.

3 Model Description

3.1 Thermal History. Detailed knowledge of the thermal history of the bars is necessary to predict the quench distortion. The average of the three recorded temperatures at 6.35 mm from the quenched surface and a one-dimensional inverse heat conduction model (Arthur, et al., 1994) based on an algorithm by Beck, et al. (1980) were employed to predict the temperature of the surface of the specimens as a function of time. Given this surface temperature history, the commercial finite element code ABAQUS (1993) was used to determine the transient temperature distribution within the bars.

3.1.1 Commercial Purity Aluminum. The temperature dependent thermal properties for the commercial purity aluminum are represented reasonably well by the expressions (Appendix A) based on data found in the literature (Hatch, 1984; Touloukian and Ho, 1973). The surface temperature history and a 100 W/m^2K heat transfer coefficient applied to the upper surface were used in the finite element model to calculate the temperatures at the thermocouple locations. The results are shown in Fig. 4. The agreement with the experiments is very good except for a short duration, 10 deg overestimate by the model at about 80 sec into the quench. The origin of this discrepancy could not be determined. The thermal history is represented reasonably well even during the rapid transient of the initial quench.

3.1.2 Al-4.6 Mg. The thermal properties located in the literature for Al-4.6 Mg were available only to 450°C (Touloukian and Ho, 1973) and were extrapolated to higher temperatures. The density and specific heat are close to that of the commercial purity aluminum given above. These same expressions were used for the Al-4.6 Mg alloy. The thermal conductivity was approximately 40 percent lower than that of the commercial purity material. The thermal conductivity data was extrapolated to 550°C using several functional forms. One that seemed to give a reasonable fit of the literature data from 20°C to 450°C is found in Appendix A. Reasonable temperature estimates at the thermocouple locations were also found.

3.2 Thermal Expansion. The temperature gradients in the specimens result in thermal strains which are the driving force for the deformation. The thermal strain is related to the temperature by

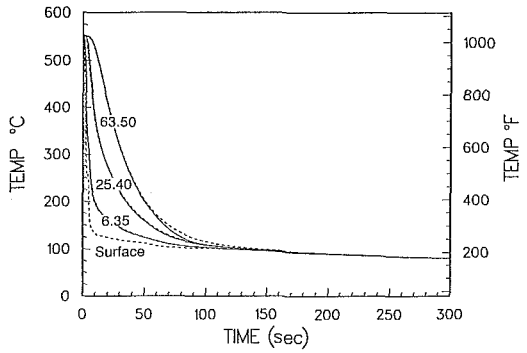


Fig. 4 Experimental curvature history for the commercial purity aluminum bar and predictions using various constitutive models

$$\boldsymbol{\epsilon}^T = [\alpha(T)(T - T_0) - \alpha(T_I)(T_I - T_0)]\mathbf{I}, \tag{2}$$

where $\boldsymbol{\epsilon}^T$ is the thermal strain tensor; $\mathbf{I}$ is the identity tensor; T_I is the initial temperature; and T_0 is some reference temperature. The thermal expansion coefficient as a function of temperature for the two alloys is given in Appendix A. The calculated thermal contraction of the bars is in very good agreement with the data.

3.3 Mechanical Properties. Deformation of the bars results from a combination of thermal strain, elastic strain, and inelastic strain. Here, an increment of total strain is assumed to be additively decomposed as

$$\Delta\boldsymbol{\epsilon} = \Delta\boldsymbol{\epsilon}^T + \Delta\boldsymbol{\epsilon}^e + \dot{\boldsymbol{\epsilon}}^{in}\Delta t. \tag{3}$$

The total strain is strictly a function of geometry, and it must satisfy compatibility. Poisson's ratio is 0.3 and the elastic modulus is given by

$$E(\text{MPa}) = 70170. - 42.62\ T, \tag{4}$$

where T is the temperature in degrees Celsius. The inelastic response of the material is dependent on the history of the deformation, so the inelastic strain rate must be integrated through time.

The inelastic response follows the normality flow rule with an isotropic, von Mises flow potential. The stress is assumed to be a function of the temperature, strain rate, and the state variables:

$$\sigma = \mathrm{F}(T, \dot{\epsilon}^{in}, S_i). \tag{5}$$

Material constitutive relations of this form were implemented as user material subroutines (Smelser and Becker, 1989) in the finite element code ABAQUS (1993).

3.3.1 Commercial Purity Aluminum. Several constitutive models have been developed for the commercial purity alloy in previous studies. The ones explored here are as follows: a two internal state variable model (TIV) (Sample et al., 1991); a single internal state variable model (SIV) (Sample and Lalli, 1987); and a hyperbolic sine model (HS), similar to the SIV model with no internal state variables. These constitutive models were all developed based on data obtained in compression tests to large strains. Data were not collected in the range of small strains pertinent in quench distortion. The models were never intended for such applications. However, for lack of more appropriate data, these existing models were used. In addition, two other models were constructed to explore the influence of variations in the constitutive model on the curvature. Details of these two constitutive models are given in Appendix B.

3.3.2 Al-4.6 Mg. The constitutive model for the Al-4.6 Mg alloy was developed from data obtained from tensile tests conducted at temperatures from 40°C to 520°C and for strain rates ranging from 10^{-6} sec^{-1} to 10^{-2} sec^{-1}. In this model the effective stress is determined by

$$\sigma = S\frac{\sigma_0}{\sigma_M}, \tag{6}$$

where S is a state variable and $\sigma_M = 100$ MPa is a constant. The flow stress for the fully annealed material, σ_0, is given by

$$\sigma_0 = \sigma_M\left[\frac{RT}{Q}\sinh^{-1}(\hat{Z}_0)\right]^m;$$

$$\hat{Z}_0 = \dot{\epsilon}^{in}\exp\left(\frac{Q}{RT} - \ln A_0\right). \tag{7}$$

The annealed flow stress depends on the effective inelastic strain rate, $\dot{\epsilon}^{in}$, and the absolute temperature, T. Here, R is the gas

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

constant; $Q = 150000$ (J/mole) the apparent activation energy; $m = 0.3$ a rate sensitivity exponent; and $\ln A_0 = 19.83$.

The state variable accounts for hardening from the annealed stress to the stress at steady-state conditions, σ_S:

$$\sigma_S = \sigma_C \sinh^{-1}\left[(\hat{Z}_S)^m\right];$$

$$\hat{Z}_S = \dot{\epsilon}^m \exp\left(\frac{Q}{RT} - \ln A_S\right). \quad (8)$$

The additional constants are $\sigma_C = 45$ MPa and $A_S = 20.29$. The initial value of the state variable is $S = \sigma_M$, and it evolves with deformation toward its steady-state value $S_S = \sigma_M \sigma_S / \sigma_0$ according to

$$\dot{S} = [H_0 - (S_S - \sigma_M)(S - \sigma_M)]\left(1 - \left(\frac{S}{S_S}\right)^2\right)\dot{\epsilon}. \quad (9)$$

$H_0 = 9600$ is a constant characterizing the hardening.

3.4 Finite Element Model. The specimens were designed such that the longitudinal stress component would be significantly larger than the other stress components over the majority of the specimen. However, the bars were truly three dimensional so the effects of the finite specimen geometry were investigated. The solutions were obtained in an uncoupled manner: The complete thermal history was calculated first, then this temperature history was imposed on the mechanical model.

A model using a single column of generalized plane-strain elements was used in the analysis. This model used 45 elements in the X_2 direction. The elements were unconstrained in the X_3 direction. In the X_1 direction the element faces are constrained to lie on two planes. The angle between these planes is used to calculate the curvature. The curvature predicted by this simpler model fell between the results from two more complicated two-dimensional analyses, whichever similar. Therefore, the results from this simple model were deemed acceptable.

4 Model Predictions

4.1 Commercial Purity Aluminum Bars. The experimental and predicted curvatures using the three constitutive models are shown in Fig. 5. Also included is the curvature prediction assuming elastic behavior. The model predictions vary widely and predict the same deformation trends. None gives quantitative agreement with the experiment. The hyperbolic sine model (the least sophisticated model) seems to give the best curvature prediction at the end of the test. However, examination of the results on a short time scale shows that predictions from this model differ more from the experiment

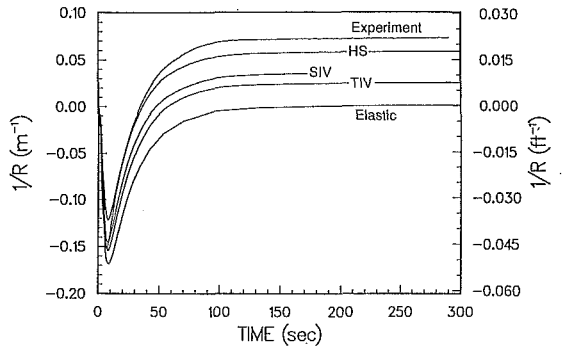


Fig. 5 Experimental curvature history on a short time scale for the commercial purity aluminum bar and predictions using various constitutive models: (HS) Hyperbolic Sine, (SIV) Single Internal Variable, (TIV) Two Internal Variable

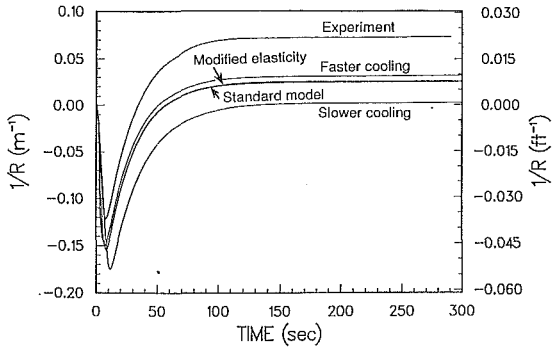


Fig. 6 Curvature predictions for the commercial purity aluminum bar with modifications to the thermal boundary conditions and thermophysical properties

during the initial quench than do the other models. Hence, the better agreement at the end of the analysis is probably coincidental.

Various modifications were made to the model in an attempt to improve the predictions. The independent changes included varying the high temperature elastic modulus by 40 percent and altering the temperature dependence of the thermal expansion coefficient by 20 percent. Changes in thermal history were also explored by modifying the imposed surface temperature history in the heat transfer portion of the analysis. A more rapid initial cooling was examined by extending the steep initial drop in the surface temperature from 295°C at 2.4 sec to 180°C at 2.8 sec. The surface temperature was then decreased linearly to merge with the original curve at 165°C at 5.5 sec. A slower surface cooling was achieved by decreasing the surface temperature linearly from the value at 1.0 sec to merge with the original curve at 9.5 sec.

Curvature predictions using some of these modifications and the TIV constitutive model are shown in Fig. 6. The changes in the elastic properties and the thermal expansion coefficient made little difference. The modifications to the thermal history had more of an effect. The accelerated cooling produced better agreement with the experiment. However, none of these changes improved the predictions substantially, particularly in the initial 10 sec of the quench.

From the results in Fig. 5, it is evident that variations in the constitutive behavior have a large effect on the curvature predictions. Hence, alternative descriptions of the material behavior were pursued. One aspect of high temperature, low strain rate deformation that is not captured by these material models is primary creep. When load is applied to a creep specimen, many materials have a high initial creep rate that slows to the steady-state creep rate level. When viewed in terms of the material behavior at a constant strain rate, this implies strain hardening at low strains.

To assess the possible effect of hardening at small strains, two constitutive models were constructed. These are given in Appendix B as the Modified Hyperbolic Sine (MHS) model and the Hardening Modified Hyperbolic Sine (HMHS) model. The MHS model is a steady-state model (no internal state variables or hardening) constructed to give approximately the same stress-strain rate-temperature relation as the TIV model at a strain of 0.01. The HMHS model provides the identical response at strains on the order of 0.01, but the flow strength is reduced at lower strains and lower temperatures. The stress-strain curves at several temperatures and strain rates are compared for the HMHS, SIV, and TIV models in Fig. 7.

Curvature predictions from simulations using these constitutive models, the experimental result, and the prediction from the TIV model are shown in Fig. 8. The MHS model gives somewhat better predictions than the TIV model. The addition of the hardening greatly improves the prediction; the deviation

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

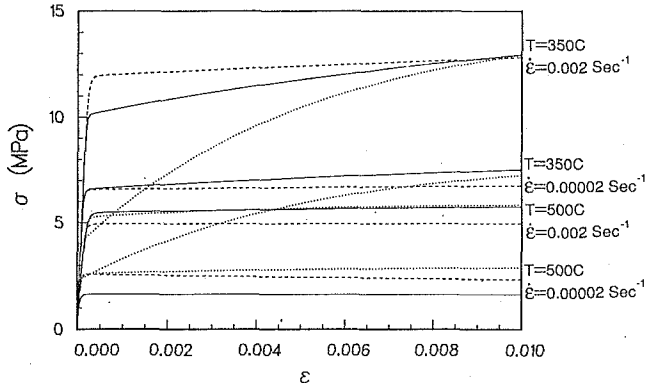


Fig. 7 Stress-strain curves at two temperatures and two strain rates for constitutive models: Two Internal Variable (TIV) (—), Single Internal Variable (SIV) (----) and Hardening Modified Hyperbolic Sine (HMHS) (· · · · · ·)

from the experiment is only half of the deviation without the hardening. This suggests that strain hardening at low strains has an impact on the quench distortion behavior.

4.2 Al-4.6 Mg Bars. The curvature prediction for the Al-4.6 Mg alloy bar is given in Fig. 9 along with the experimental result. There is good agreement on the initial bending, and reasonable agreement for the remainder of the experiment. As with the commercial purity material, the deviations could have been introduced in modelling the thermal boundary conditions, representing the thermophysical properties or the constitutive behavior. The effects of variation of these parameters were explored above and will not be examined further.

The predicted longitudinal residual stress in the Al-4.6 Mg bar is given as the solid line in Fig. 10. The quenched surface is at zero on the abscissa. The knee in the curve near the quenched surface results from plastic flow under compressive stress toward the end of the analysis. The agreement with the experiment is very good over most of the specimen. There is, however, a significant discrepancy near the quenched surface. This will be addressed below.

5 Discussion

The experiment presented here is conceptually simple. There is one dominant stress component, and the temperature and stress gradients are significant in only one direction. In contrast to this simplicity, the details of the response of the bars are complex. Upon initial quenching, the material near the quenched surface contracts causing the bar to arch away from the water. During this time, both the quenched surface and the opposite surface are in tension and the material elongates inelastically in these regions. When the lower surface of the

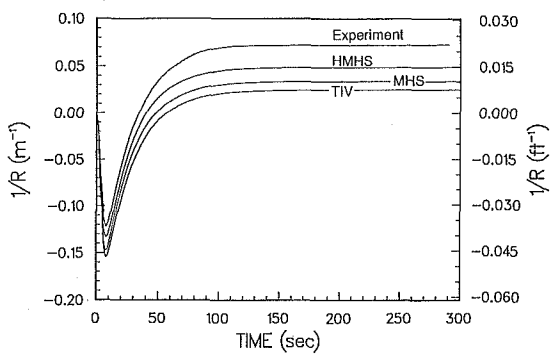


Fig. 8 Curvature predictions for the commercial purity aluminum bar using the modified constitutive relations

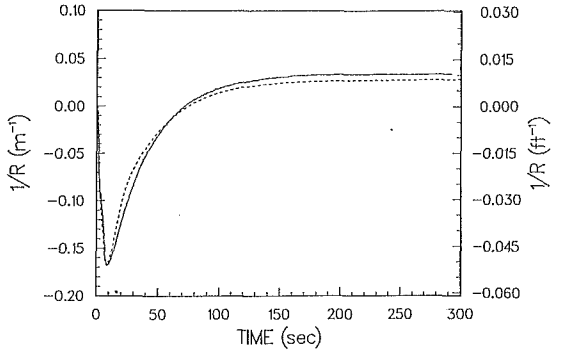


Fig. 9 Measured (—) and predicted (----) curvature histories for the Al-4.6 Mg alloy bar

bar cools sufficiently, the rate of contraction on the quenched surface is equal to the rate of contraction on the opposite surface, and the rate of change of curvature is zero. This corresponds to the minimum curvature at approximately 9 sec in Fig. 3. As the cooling continues, the magnitude of the curvature decreases. At some later point the longitudinal stresses at the surfaces become compressive since the material at the surfaces has elongated inelastically during the early part of the quench. This reverses the direction of inelastic straining.

The final curvature depends on the ratio of the time-integrated inelastic strains on the two surfaces. If the total inelastic stretching is greater on the quenched surface, the final configuration will have the center of the bar bowed down into the water. This is the case in the present experiments. This simple explanation of the deformation can serve as a conceptual model to envision how the curvature would be altered by changing the bar height, quench rate, and geometric constraints.

Predictions of quench distortion and stress development are also complicated by details of the material behavior. The material experiences at least one stress reversal, and the constitutive model should be sophisticated enough to treat any flow stress reductions on reverse loading. The models used here did not. Anelasticity should also be included. Additional complications arise for heat treatable alloys in which significant precipitation of alloying elements can occur over a short time period. For such materials the thermal and mechanical properties change not only as a function of temperature but as a function of time due to solute loss and precipitate formation. The difficulties are compounded by possible microstructural gradients across the specimens resulting from macrosegregation and nonuniform thermomechanical processing. Many of these complications were avoided here by selecting materials with minimal precipitation strengthening, no phase transformations and little grain size dependence.

The present analyses demonstrate that good quantitative predictions require accurate descriptions of the thermal boundary conditions, the thermophysical properties and the material constitutive behavior. Complete property and process data are essential ingredients in the modeling. The property data used here, however, were shown to be inadequate. In this study, modeling of the thermal boundary conditions was avoided by collecting transient temperature data as a basis for the thermal solution.

One of the most troubling results from this study was the residual stress measurement at the surface of the bar. The stress estimated from a standard layer removal procedure was approximately a factor of two greater than the room temperature flow strength, Fig. 10. A high measured residual stress could result from several factors, but few would explain a stress twice as high as the yield stress. The possibility of a high strength gradient near the surface was explored by machining a thin sheet-like tensile specimen from the surface of one of the quenched bars. The yield strength in this sample was approximately 85

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

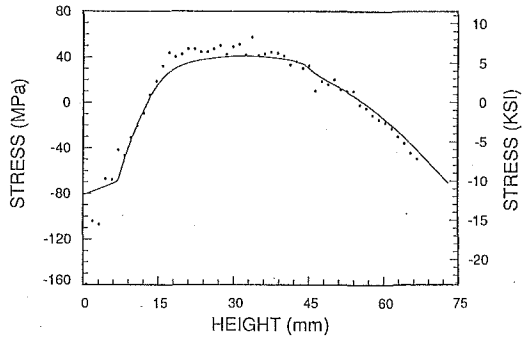


Fig. 10 Predicted (—) and measured (●) longitudinal residual stress for the Al-4.6 Mg alloy bar

MPa, the same as given by the constitutive model. Experimental errors are always a possibility, but in order to err by a factor of two, substantial errors in the strain gage measurement or in the depth of the layer removed would be needed. Since there are other data points near 105 MPa, the data are consistent, and this seems unlikely. A more likely experimental error is that the clamping or machining operations induced plastic deformation in the bar during the layer removal procedure. However, this hypothesis has not been confirmed.

From Figs. 9 and 10 it appears that the residual stress predictions are in better agreement with experiments than the curvature predictions. In these simple bar experiments, the stress is governed largely by the temperature field and by force and moment equilibrium. The stress is governed to a lesser extent by the strain rate through the constitutive model. The deformation, on the other hand, has no external constraints and is determined largely by the constitutive relation. The greater error in the distortion predictions can be explained by considering the nonlinearity of the constitutive model. Suppose that at some point in the analysis the stress in the model is five percent in error. A five percent stress variation in the constitutive model gives an inelastic strain rate that can easily vary by 50 percent, depending on the stress and temperature. Thus, for particular geometries, the residual stress predictions can be close even if the distortion error is significant. This is consistent with the results from Fletcher and Soomro (1986) who examined the effect of martensite start and finish temperatures on residual stress in plate. Their calculations show little variation in residual stress predictions for significant differences in residual strain predictions.

The seemingly good quench distortion predictions for the Al-4.6 Mg alloy bar may be fortuitous as a result of little plastic strain in these tests. Additional tests with greater plastic strain are needed to validate the predictive capabilities of the constitutive model.

The extreme sensitivity of quench distortion in these experiments to changes in material properties and thermal boundary conditions may not be found in other shapes where geometric constraints limit distortion. Additional investigations on geometries with varying degrees of constraint should be conducted to determine how sensitive quench distortion and residual stress are to the thermal history and to the material constitutive behavior.

Acknowledgments

The authors would like to thank W. D. Bennon for support of the program and H. E. Clements, G. P. Nivin, and R. W. Schultz for their assistance in the experiments.

References

ABAQUS, 1993, Hibbitt, Karlsson & Sorensen, Inc., Pawtucket, RI.

Aksel, B., 1990, "Analysis of Stress and Displacement Fields Generated During Quenching and Casting," Ph.D. Dissertation, Cornell University, Ithaca, NY.

Aksel, B., Arthur, W. R., and Mukherjee, S., 1992, "A Study of Quenching: Experiment and Modeling," *Journal of Engineering for Industry,* Vol. 114, pp. 309–316.

Arthur, W. R., Becker, R. C., and Karabin, M. E., 1994, "Experimentation and Modeling of Residual Stresses and Distortion in Quenched Rectangular Bars," Proceedings of the International Heat Treating Conference, *Heat Treating: Equipment and Processes,* G. E. Totten and R. A. Wallis, eds ASM International, Materials Park, OH, pp. 43–50.

Beck, J. V., Litkouhi, B., and St. Clair, Jr., C. R., 1980, "Efficient Numerical Solution of the Nonlinear Inverse Heat Conduction Problem," ASME Paper No. 80-HT-3.

Becker, R., Karabin, M. E., Liu, J. C., and Smelser, R. E., 1994, "Experimental Validation of Predicted Distortion and Residual Stress in Quenched Bars," *Computational Material Modeling,* eds A. K. Noor and A. Needleman, eds. ASME, New York.

Boyer, J. C., and Boivin, M., 1985, "Calculations of Residual-Stress Relaxation in Quenched Plates," *Material Science and Technology,* Vol. 1, pp. 786–792.

Fletcher, A. J., and Price, R. F., 1981, "Generation of Thermal Stress and Strain During Quenching of Low-Alloy Steel Plates," *Metals Technology,* Nov., pp. 427–446.

Fletcher, A. J., and Soomro, A. B., 1986, "Effect of Transformation Temperature Range on Generation of Thermal Stress and Strain During Quenching," *Material Science and Technology,* Vol. 2, pp. 714–719.

Hatch, J. E. ed., 1984, *Aluminum: Properties and Physical Metallurgy,* American Society for Metals, Metals Park, OH.

Jeanmart, P., and Bouvaist, J., 1984, "Finite Element Calculation and Measurement of Thermal Stresses in Quenched Plates of High Strength 7075 Aluminum Alloy," *Calculation of Internal Stresses in Heat Treatment of Metalic Materials,* Linköping University, Sweden, pp. 23–42.

Landau, H. G., and Weiner, J. H., 1958, "Transient and Residual Stresses in Heat-Treated Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 459–465.

Sample, V. M., and Lalli, L. A., 1987, "Effects of Thermomechanical History on Hardness of Aluminum," *Materials Science and Technology,* Vol. 3, pp. 28–35.

Sample, V. M., Lalli, L. A., and Richmond, O., 1991, "Phenomenological Concepts for the Development of a Two Internal State Variable Model," *Modeling the Deformation of Crystalline Solids,* T. C. Lowe, A. D. Rollett, P. S. Follansbee, and G. S. Daehn, eds., The Materials, Metals & Materials Society, Warrendale, PA, pp. 327–345.

Smelser, R. E., and Becker, R., 1989, "ABAQUS User Subroutines for Material Modeling," *Proceedings of ABAQUS Users' Conference,* Strasa, Italy, Hibbitt, Karlsson & Sorensen, Providence, RI, pp. 207–226.

Touloukian, Y. S., and Ho, C. Y., eds., 1973, *Properties of Aluminum and Aluminum Alloys,* Purdue University, West Lafayette, IN.

APPENDIX A

Thermal Properties

A1 Commercial Purity Aluminum. The temperature-dependent thermal properties for the commercial purity aluminum found in Hatch (1984) and Touloukian and Ho (1973) are represented reasonably well by

$$\rho(\text{g/cm}^3) = 2.7889 - 0.2152 \times 10^{-3}\, T$$

$$C_p(\text{J/g K}) = 0.7639 + 0.4644 \times 10^{-3}\, T$$

$$K(\text{W/mm K}) = 0.2473 - 4.52 \times 10^{-6}\, T - 40.2 \times 10^{-9}\, T^2. \quad (A1)$$

Here, ρ is the density; C_p is the specific heat; K is the thermal conductivity; and T is the temperature in Kelvin. The thermal expansion coefficient, $\alpha(T)$, was fit as

$$\alpha(T) = 22.23 \times 10^{-6} + 9.78 \times 10^{-9}\, T, \quad (A2)$$

where T is the Celsius temperature.

A2 Al-4.6 Mg. A reasonable fit of the literature data from 20°C to 450°C for the thermal conductivity data was

$$K(\text{W/mm K}) = 0.01\, T^{0.45}, \quad (A3)$$

where T is the temperature in Kelvin. The thermal expansion data for Al-4.6 Mg was only available to 450°C. The expansion data was extrapolated to 550°C by several different curve fits. The estimated thermal expansion coefficient for this alloy is

$$\alpha(T) = 22.23 \times 10^{-6} + 9.78 \times 10^{-9}\, T$$

for $20°\text{C} < T < 400°\text{C}$;

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\alpha(T) = 11.95 \times 10^{-6} + 35.73 \times 10^{-9}\, T$$

$$\text{for } 400°\text{C} < T < 550°\text{C}. \quad (A4)$$

APPENDIX B

Material Constitutive Models

B1 MHS Constitutive Model. The Modified Hyperbolic Sine (MHS) model is an ad hoc steady-state stress model that gives approximately the same stresses as the TIV model at 1 percent strain. It gives reasonable initial flow strengths at room temperature, unlike the steady-state HS model. The flow stress for the MHS model is given by

$$\sigma(\text{MPa}) = \sigma_{ss} = 12.0 \sinh^{-1}(\bar{Z}^{1/6.25});$$

$$\bar{Z} = \dot{\epsilon}^{in} \exp\left(\frac{266070}{R(T + 150)} - 32.51\right). \quad (B1)$$

$R = 8.314$ J/mol-K is the gas constant and T is the absolute temperature in Kelvin.

B2 HMHS Constitutive Model. The HMHS model incorporates a temperature-dependent strain hardening into the MHS model. The flow stress is obtained by multiplying the stress from the MHS model by

$$1 - \exp\left(-\frac{\epsilon^{in}}{0.005}\right) \exp\left(-\left[\frac{(T - 373)}{335}\right]^4\right), \quad (B2)$$

where ϵ^{in} is the inelastic strain and temperature is in Kelvin.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Micromechanical Definition of the Strain Tensor for Granular Materials

N. P. Kruyt
Department of Mechanical Engineering,
University of Twente,
Enschede, The Netherlands

L. Rothenburg
Department of Civil Engineering,
University of Waterloo,
Waterloo, Ontario,
Canada

In order to develop constitutive relations for granular materials from the micromechanical viewpoint, general expressions relating macroscopic stress and strain to contact forces and particle displacements are required. Such an expression for the stress tensor under quasi-static conditions is well established in the literature, but a corresponding expression for the strain tensor has been lacking so far. This paper presents such an expression for two-dimensional assemblies. This expression is verified by computer simulations of biaxial and shear tests. As a demonstration of the use of the developed expression, a study is made of the elastic moduli of two-dimensional, isotropic assemblies of bonded, nonrotating disks. Theoretical expressions are given for the elastic moduli in terms of micromechanical parameters, such as coordination number and contact stiffnesses. Comparison with the results from computer simulations show that the agreement is fairly good over a wide range of coordination numbers and contact stiffness ratios.

Introduction

Constitutive relations describing the behavior of granular materials are of great importance to various geotechnical and industrial applications. Usually these constitutive relations are developed from the continuum-mechanical viewpoint and do not recognize the discrete nature of granular materials. The resulting relations are frequently phenomenological in nature.

As an alternative to the continuum-mechanical approach, the micromechanical approach to constitutive modeling of granular materials under quasi-static conditions is being developed (for example, Cundall et al., 1982; Bathurst and Rothenburg, 1988a, b; Rothenburg and Bathurst, 1989; Rothenburg et al., 1989; Mehrabadi et al., 1993). Herein a granular material is modeled as an assembly of semi-rigid particles interacting by means of contact forces. Development of constitutive relations is performed using suitable averaging techniques.

In order to link the behavior on the micro (particle) level to the macro (continuum) level, general micromechanical expressions for the stress and strain tensors are required. The expression for the stress tensor is well established (Drescher and De Josselin de Jong, 1972; Strack and Cundall, 1978; Rothenburg and Selvadurai, 1981), but a similar expression for the strain tensor has been lacking so far in the literature. The principal aim of this paper is to develop such a general expression for the two-dimensional case. This expression will be verified by computer simulation of biaxial and direct shear tests.

As a demonstration of the use of the resulting micromechanical expressions for the stress and strain tensor, a study is made of the elastic moduli of two-dimensional, isotropic assemblies of bonded, nonrotating disks. Theoretical expressions will be derived for the elastic moduli in terms of micromechanical parameters, such as contact stiffnesses and coordination number, i.e., the average number of contacts per particle. These theoretical expressions will be compared with the results of computer simulations.

The usual sign convention from continuum mechanics for stress and strain is adopted: tensile stresses and strains are counted positive. The summation convention is adopted: summation is implied over repeated indices.

Micromechanical Stress Tensor

Many authors (Drescher and De Josselin de Jong, 1972; Strack and Cundall, 1978; Rothenburg and Selvadurai, 1981) have proposed an expression for the average stress tensor in terms of the forces acting at the contacts between particles and the geometry of the assembly of particles. A "derivation" of this expression is repeated here, since it suggest an analogous way to "derive" the expression for the average strain tensor.

The derivation of the expression for the average stress tensor proceeds in two steps. In the first step the average stress tensor is related to quantities involving forces exerted on the particles by the boundary that encloses the assembly of particles. The second step equates these quantities involving external forces to quantities involving internal forces. The result is the micromechanical expression for the average stress tensor.

Average Stress Tensor in Terms Involving External Forces. The expression for the average stress tensor is derived under conditions of quasi-static equilibrium and in the absence of body forces. Then the (continuum) equilibrium conditions are

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0. \tag{1}$$

The two-dimensional average stress tensor in area S with boundary B is defined by

$$\bar{\sigma}_{ij} = \frac{1}{S}\int_S \sigma_{ij} dS. \tag{2}$$

From (1) and Gauss' theorem it follows

$$\bar{\sigma}_{ij} = \frac{1}{S}\int_B n_k \sigma_{ik} x_j dS \tag{3}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechancial Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Feb. 11, 1994; final revision, June 15, 1995. Associate Technical Editor: J. T. Jenkins.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where n_i is the vector normal to the boundary; the vector tangential to the boundary is denoted by t_i. Considering the loads to be point loads on the boundary B, it follows:

$$\bar{\sigma}_{ij} = \frac{1}{S} \sum_{\beta \in B} f_i^\beta x_j^\beta \tag{4}$$

where f_i^β is the boundary force exerted on boundary particle β. This is the expression for the average stress tensor in terms involving external forces.

Average Stress Tensor in Terms Involving Internal Forces. The equilibrium conditions for particle p in the absence of body forces read

$$\sum_q f_i^{pq} = 0 \tag{5}$$

where the summation is over the particles q that are in contact with particle p and f_i^{pq} is the force exerted by particle q on particle p.

Multiplication of (5) by the position vector X_j^p of the center of mass of particle p and addition of all equations gives

$$\sum_p \sum_q f_i^{pq} X_j^p = 0. \tag{6}$$

This double sum contains one term for each boundary contact β with particle p $f_i^\beta X_j^p$, which can be rewritten as $f_i^\beta (x_j^\beta - l_j^{p\beta})$, where $l_j^{p\beta}$ is the so-called *contact vector* connecting the center of mass of particle p to the boundary contact point β.

Each internal contact between particles p and q contributes a term $(f_i^{pq} X_j^p + f_i^{qp} X_j^q)$. Since $f_i^{pq} = -f_i^{qp}$, terms corresponding to internal contacts can be written as $-f_i^{pq}(X_j^q - X_j^p)$ or $-f_i^{pq} l_j^{pq}$, where l_j^{pq} is the so-called *contact vector* connecting the centres of the particles p and q. Combinations $f_i^{pq} l_j^{pq}$ can be written as $f_i^c l_j^c$, since $f_i^{pq} = -f_i^{qp}$ and $l_i^{pq} = -l_i^{qp}$. As a result it follows that

$$\sum_{\beta \in B} f_i^\beta x_j^\beta = \sum_{c \in S} f_i^c l_j^c. \tag{7}$$

Hence it follows from (4),

$$\bar{\sigma}_{ij} = \frac{1}{S} \sum_{c \in S} f_i^c l_j^c. \tag{8}$$

This is the expression for the average stress tensor in terms involving internal forces, i.e., the micromechanical expression for the average stress tensor.

Average Stress Tensor in Terms of Group Averages. Bathurst and Rothenburg (1988a, b) have suggested grouping the contacts within a finite number of orientation classes. Then group averages $\overline{f_i^c l_j^c}(\varphi_g)$ can be calculated and relation (8) can be rewritten in terms of group averages as

$$\bar{\sigma}_{ij} = \frac{M_S}{S} \sum_g \overline{f_i^c l_j^c}(\varphi_g) E(\varphi_g) \Delta\varphi \tag{9}$$

where M_S is the total number of contacts in area S and $E(\varphi)$ is the contact orientation distribution function, as proposed by Horne (1965). $E(\varphi)\Delta\varphi$ is the fraction of contacts with orientations within $(\varphi, \varphi + \Delta\varphi)$.

The continuous form of (9), valid for an infinite assembly, becomes

$$\bar{\sigma}_{ij} = m_S \int_0^{2\pi} \overline{f_i^c l_j^c}(\varphi) E(\varphi) d\varphi \tag{10}$$

where $m_S = M_S/S$ is the contact density with respect to assembly area.

Micromechanical Strain Tensor

In analogy to the expression for the average stress tensor, an expression for the average displacement gradient is derived here. The strain tensor is obtained by taking the symmetric part of the displacement gradient tensor.

The derivation of the expression for the average displacement gradient tensor proceeds in two steps. In the first step the average displacement gradient tensor is related to quantities involving relative displacements of the boundary particles. The second step equates these quantities involving external relative displacements to quantities involving internal relative displacements. The result is the micromechanical expression for the average displacement gradient tensor, and hence for the average strain tensor.

Average Displacement Gradient Tensor in Terms Involving External Relative Displacements. The average displacement gradient tensor is defined by

$$\bar{\theta}_{ij} = \frac{1}{S} \int_S \frac{\partial u_i}{\partial x_j} dS \tag{11}$$

where u_i is the displacement vector. Using Gauss' theorem it follows that

$$\bar{\theta}_{ij} = \frac{1}{S} \int_B u_i n_j ds. \tag{12}$$

A relation based on (12) has been proposed by Strack and Cundall (1978). Constitutive relations at the contact will involve *relative* displacements between particles. Therefore it is desirable to transform (12) to a form containing derivatives of the displacements. This is done using the following identity:

$$\int_B u_i t_k ds = \int_B u_i \frac{dx_k}{ds} ds = -\int_B \frac{du_i}{ds} x_k ds. \tag{13}$$

Combining (12) and (13) gives

$$\bar{\theta}_{ij} = -\frac{1}{S} e_{jk} \int_B \frac{du_i}{ds} x_k ds \tag{14}$$

where e_{ij} is the two-dimensional permutation tensor. The discrete formulation of (14) in terms of relative displacements at the boundary is

$$\bar{\theta}_{ij} = -\frac{1}{S} e_{jk} \sum_{\alpha \in B} \Delta l_i^\alpha x_k^\alpha. \tag{15}$$

This expression for the average displacement gradient tensor is analogous to Eq. (4) for the average stress tensor. It gives the average displacement gradient tensor in terms involving external relative displacements.

Average Displacement Gradient Tensor in Terms Involving Internal Relative Displacements. The derivation of the stress tensor employed the equilibrium conditions for the particles. The equivalents for the displacement gradient tensor are the compatibility conditions for polygons. These polygons arise as a way of dividing the plane network of particle centers of mass and contacts into polygons, as depicted in Fig. 1. Various properties associated with such a subdivision of the assembly into polygons were studied by Satake (1992).

Since the polygons form closed loops, the compatibility conditions for polygon r are

$$\sum_s \Delta l_i^{rs} = 0 \tag{16}$$

where the summation is over the sides of polygon r and Δl_i^{rs} is the relative displacement between particles comprising side s of polygon r. Multiplication of (16) by the position vector

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

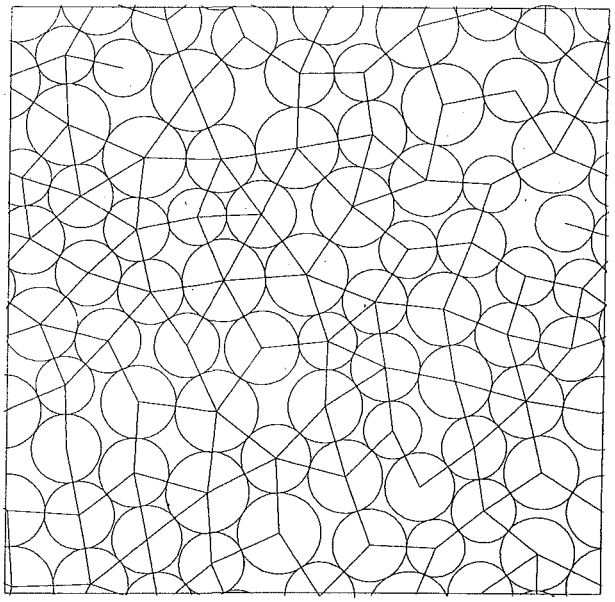


Fig. 1 Tessellation of the area into polygons, based on the contacts of the particles

V_i^r of the center of gravity of polygon r and addition of all equations gives

$$\sum_r \sum_s \Delta l_i^{rs} V_j^r = 0. \tag{17}$$

This double sum contains one term for each external side α of polygon r $\Delta l_i^{r\alpha} V_j^r$, which can be rewritten as $\Delta l_i^\alpha (x_j^\alpha - g_j^{r\alpha})$, where $g_j^{r\alpha}$ is the vector connecting the center of gravity of polygon r to boundary point α.

Each internal side contributes a term $(\Delta l_i^{rs} V_j^r + \Delta l_i^{sr} V_j^s)$. Since $\Delta l_i^{rs} = -\Delta l_i^{sr}$, terms corresponding to internal contacts can be written as $-\Delta l_i^{rs}(V_j^s - V_j^r)$ or $-\Delta l_i^{rs} g_j^{rs}$, where g_j^{rs} is the vector connecting the centers of gravity of polygons r and s. Combinations $\Delta l_i^{rs} g_j^{rs}$ can be written as $\Delta l_i^c g_j^c$, since $\Delta l_i^{rs} = -\Delta l_i^{sr}$, and $g_j^{rs} = -g_j^{sr}$. The resulting expression for (17) becomes

$$\sum_{\alpha \in B} \Delta l_i^\alpha x_j^\alpha = \sum_{c \in S} \Delta l_i^c g_j^c. \tag{18}$$

Hence it follows from (15)

$$\bar{\theta}_{ij} = \frac{1}{S} \sum_{c \in S} \Delta l_i^c h_j^c \tag{19}$$

where the so-called *polygon vector* h_j^c is defined by

$$h_j^c = -e_{jk} g_k^c. \tag{20}$$

Equation (19) is the expression for the average displacement gradient tensor in terms involving internal relative displacements. This micromechanical expression for the average displacement gradient tensor is analogous to the micromechanical expression for the average stress tensor (8). Equation (19) was first reported by Rothenburg (1980).

The expression for the average strain tensor then becomes

$$\bar{\epsilon}_{ij} = \frac{1}{S} \sum_{c \in S} \frac{1}{2} (\Delta l_i^c h_j^c + \Delta l_j^c h_i^c). \tag{21}$$

Average Strain Tensor in Terms of Group Averages. In analogy to Eq. (10) for the average stress tensor, the average strain tensor can be expressed in terms of group averages. The continuous form for the average strain tensor, valid for an infinite assembly, is

$$\bar{\epsilon}_{ij} = m_S \int_0^{2\pi} \frac{1}{2} \overline{\Delta l_i^c h_j^c + \Delta l_j^c h_i^c}(\varphi) E(\varphi) d\varphi. \tag{22}$$

Geometrical Relation

A useful geometrical relation is derived by repeating the derivation leading to (18), but with Δl_i^c replaced by l_i^c in (17). It follows that

$$\sum_{\alpha \in B} l_i^\alpha x_j^\alpha = \sum_{c \in S} l_i^c g_j^c. \tag{23}$$

From Gauss' theorem for the area of S it follows after some algebra that

$$S\delta_{ij} = -e_{jk} \sum_{\alpha \in B} l_i^\alpha x_k^\alpha, \tag{24}$$

and hence

$$\delta_{ij} = \frac{1}{S} \sum_{c \in S} l_i^c h_j^c \tag{25}$$

where δ_{ij} is the Kronecker symbol. This means that l_i^c and h_i^c are colinear on average. In terms of group averages, the continuous form of (25) is

$$\delta_{ij} = m_S \int_0^{2\pi} \overline{l_i^c h_j^c}(\varphi) E(\varphi) d\varphi. \tag{26}$$

Uniform strain and stress can now be characterized by, respectively,

$$\Delta l_i^c = \bar{\epsilon}_{ij} l_j^c \tag{27}$$

$$f_i^c = \bar{\sigma}_{ij} h_j^c \tag{28}$$

as can be verified from (8), (19), and (25).

Verification of the Micromechanical Strain Definition

The developed micromechanical strain definition for granular materials is verified by computer simulation of a biaxial test and a shear test. The average strain tensor according to the micromechanical definition is compared with the macroscopic strain tensor determined from the displacement of boundaries of the assembly.

The computer simulation is performed using the discrete element method as proposed by Cundall and Strack (1979). This essentially is an explicit time-stepping scheme for solving Newton's equations of motion. The constitutive relation at the contact as employed here is identical to that of Cundall and Strack (1979). It involves linear springs in normal and tangential direction. The tangential force is limited by dry Coulomb friction. The assembly consists of 1001 disks with various disk radii. In the biaxial test the assembly is contained in a box with height H and width B. Initial height and width are denoted by H_0 and B_0. The macroscopic (logarithmic) displacement gradient is determined from the geometry of the box by

$$\bar{\theta}_{ij} = \begin{bmatrix} \log \dfrac{B}{B_0} & 0 \\ 0 & \log \dfrac{H}{H_0} \end{bmatrix}. \tag{29}$$

In the constant volume shear test the assembly is contained in a parallelogram with angle α, constant height H, and constant

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

width B. The macroscopic displacement gradient is determined from the geometry of the parallelogram by

$$\bar{\theta}_{ij} = \begin{bmatrix} 0 & \tan\alpha \\ 0 & 0 \end{bmatrix}. \tag{30}$$

For both the shear test and the biaxial test the macroscopic and microscopic displacements gradients were virtually identical: the differences were of the order of the roundoff error. This result completes the verification of the micromechanical strain definition.

Elastic Moduli of Isotropic Assemblies of Bonded, Nonrotating Particles

The micromechanical expressions for the stress and strain tensors are used here to derive relations for the elastic moduli of two-dimensional isotropic systems of bonded, *nonrotating* disks. This work is an extension of Bathurst and Rothenburg (1988a, b) to a wider range of coordination numbers, but restricted to the case of nonrotating particles.

The link between contact forces and contact displacements is made through the constitutive relation at the contact. A simple linear contact model without particle rotation is assumed:

$$f_n^c = k_n \Delta l_n^c \tag{31}$$

$$f_t^c = k_t \Delta l_t^c \tag{32}$$

where f_n^c and f_t^c are the normal and tangential forces at the contact and Δl_n^c and Δl_t^c are the normal and tangential relative displacements at the contact. Parameters k_n and k_t are the normal and tangential contact stiffnesses. The ratio of tangential over normal stiffness k_t/k_n is denoted by λ.

In the isotropic case considered, bulk modulus K and the shear modulus G relate macroscopic stress and strain by

$$\sigma_{11} + \sigma_{22} = 2K(\epsilon_{11} + \epsilon_{22}) \tag{33}$$

$$\sigma_{11} - \sigma_{22} = 2G(\epsilon_{11} - \epsilon_{22}) \tag{34}$$

$$\sigma_{12} = 2G\epsilon_{12}. \tag{35}$$

Geometrical Considerations. Isotropy of the contact distribution means that

$$E(\varphi) = \frac{1}{2\pi}. \tag{36}$$

Contact and polygon vectors are (approximately) related to orientation by, respectively,

$$\overline{l_i^c}(\varphi) \cong l_0 n_i(\varphi) \tag{37}$$

$$\overline{h_i^c}(\varphi) \cong h_0 n_i(\varphi) \tag{38}$$

where l_0 and h_0 are the average lengths of the contact and polygon vectors.

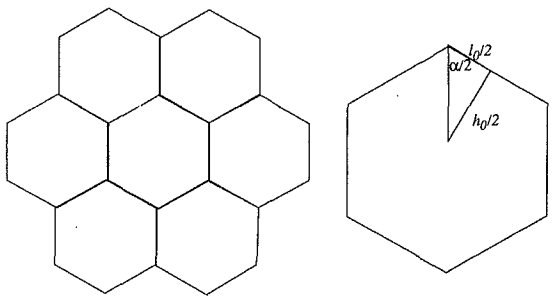


Fig. 2 (*a*) Polygons for a regular packing, (*b*) geometry of a single polygon

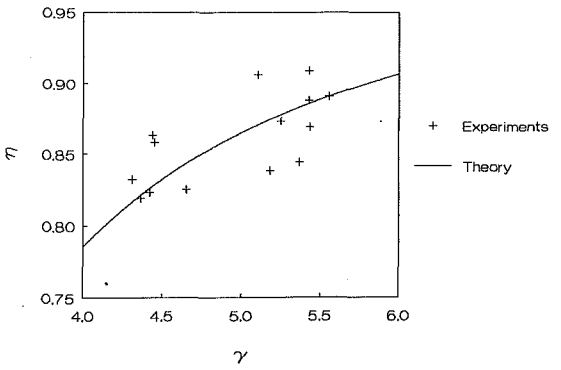


Fig. 3 Comparison of the theoretical relation for the packing density with experimental results of Quickenden and Tan (1974)

From the geometrical identity (26) it follows that

$$m_s l_0 h_0 \equiv 2. \tag{39}$$

A simple approximate relation for the ratio of the lengths of the polygon and contact vector l_0/h_0 is derived here. Assume that the assembly is regular with coordination number γ (see Fig. 2). The average internal angle α between two sides of a polygon is then given by

$$\alpha = \frac{2\pi}{\gamma}. \tag{40}$$

Considering the triangle formed by the center of the polygon, the midpoint of a side and a vertex it follows that

$$\tan\frac{\alpha}{2} = \frac{h_0}{l_0} \tag{41}$$

and hence

$$\frac{h_0}{l_0} = \tan\frac{\pi}{\gamma}. \tag{42}$$

For assemblies consisting of equal-sized disks an approximate relation for the packing density η, i.e., the fraction of area filled by the disks, can be derived with (39) and (42):

$$\eta \cong \frac{\pi}{\gamma \tan\dfrac{\pi}{\gamma}}. \tag{43}$$

In Fig. 3 this theoretical expression is compared with experimental results from (Quickenden and Tan, 1974). The agreement is acceptable.

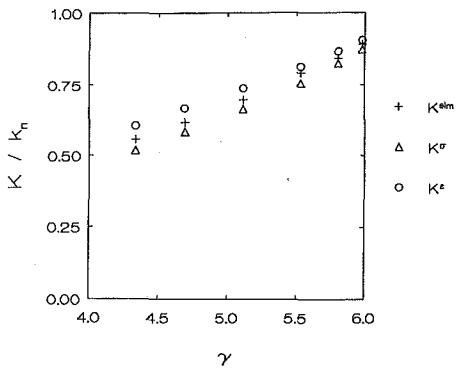


Fig. 4 Comparison between bulk moduli based on the uniform stress and strain assumptions and bulk modulus as determined from computer simulations for various coordination numbers and stiffness ratio $\lambda = 0.5$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Uniform Strain. Uniform strain is characterized by (27). Application of the constitutive relation at the contact gives

$$\sigma_{ij} = A_{ijkl}\epsilon_{kl} \tag{44}$$

where

$$A_{ijkl} = m_S l_0^2 k_n \int_0^{2\pi} (n_i n_j n_k n_l + \lambda t_i n_j t_k n_l) E(\varphi) d\varphi. \tag{45}$$

Evaluating the integrals gives the bulk modulus K^ϵ and the shear modulus G^ϵ, based on the uniform strain assumption

$$K^\epsilon = \frac{m_S l_0^2 k_n}{4} \tag{46}$$

$$G^\epsilon = \frac{m_S l_0^2 k_n}{4} \frac{1+\lambda}{2}. \tag{47}$$

Uniform Stress. Uniform stress is characterized by (28). Application of the constitutive relation at the contact gives

$$\epsilon_{ij} = B_{ijkl}\sigma_{kl} \tag{48}$$

where

$$B_{ijkl} = \frac{m_S h_0^2}{k_n} \int_0^{2\pi} \left(n_i n_j n_k n_l + \frac{1}{2\lambda} (t_i n_j + t_j n_i) t_k n_l \right) E(\varphi) d\varphi. \tag{49}$$

Evaluating the integrals gives the bulk modulus K^σ and the shear modulus G^σ, based on the uniform stress assumption

$$K^\sigma = \frac{k_n}{m_S h_0^2} \tag{50}$$

$$G^\sigma = \frac{k_n}{m_S h_0^2} \frac{2\lambda}{1+\lambda} \tag{51}$$

General Case: A Heuristic Argument. Based on the assumption of uniform strain and stress, two expressions have been derived for the bulk and shear modulus. Here a heuristic argument is given to propose expressions for the general case. Consider the case $\lambda = 0$. Conditions of static equilibrium require that $4 \leq \gamma$, conditions of compatibility for the polygons require that $\gamma \leq 6$. Hence it is expected that the expressions for the bulk and shear modulus based on the *uniform stress* assumption are valid for coordination number γ around 4 and that the expressions for the bulk and shear modulus based on the *uniform strain* assumption are valid for coordination number γ around 6.

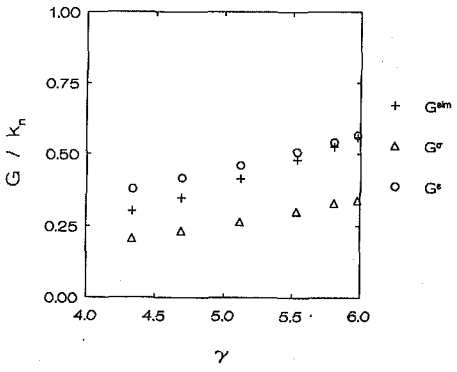


Fig. 5 Comparison between shear moduli based on the uniform stress and strain assumptions and shear modulus as determined from computer simulations for various coordination numbers and stiffness ratio $\lambda = 0.5$

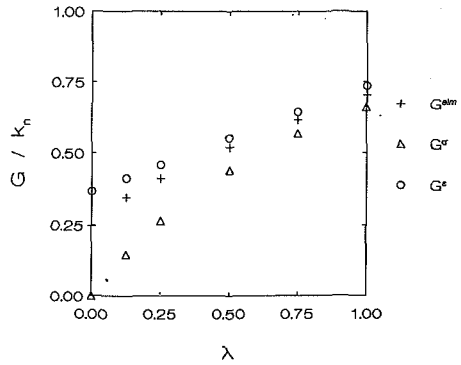


Fig. 6 Comparison between shear moduli based on the uniform stress and strain assumptions and shear modulus as determined from computer simulations for various stiffness ratios and coordination number $\gamma = 5.12$

The general heuristic expression for the bulk and shear modulus proposed here is obtained by (weighted) interpolation of the moduli obtained from the uniform strain and stress assumption:

$$K(\gamma) = \frac{\gamma - 4}{2} K^\epsilon + \frac{6-\gamma}{2} K^\sigma \tag{52}$$

$$G(\gamma) = \frac{\gamma - 4}{2} G^\epsilon + \frac{6-\gamma}{2} G^\sigma. \tag{53}$$

Using the geometrical relations (39) and (42) it follows that

$$K(\gamma) = \frac{\gamma - 4}{2} \frac{k_n}{2}\sqrt{3} + \frac{6-\gamma}{2} \frac{k_n}{2} \tag{54}$$

$$G(\gamma) = \frac{\gamma - 4}{2} \frac{k_n}{2}\sqrt{3} \frac{1+\lambda}{2} + \frac{6-\gamma}{2} \frac{k_n}{2} \frac{2\lambda}{1+\lambda}. \tag{55}$$

Computer Simulations. Computer simulations of two-dimensional isotropic assemblies of bonded, nonrotating disks have been performed using the discrete element method. A relatively narrow particle size distribution was employed for the disk radii. Simulations were performed over the range of coordination numbers $4 \leq \gamma \leq 6$ and stiffness ratios $0 \leq \lambda \leq 1$.

Figure 4 gives the comparison between the bulk moduli based on the uniform stress and strain assumptions and the bulk modulus as determined from the computer simulations for stiffness ratio $\lambda = 0.5$. A similar comparison for the shear moduli is presented in Fig. 5. For $\gamma = 5.12$, the shear moduli based on the uniform stress and strain assumptions are compared with the shear modulus as determined from the computer simulations in Fig. 6. These figures seem to indicate that the elastic moduli based on the uniform strain assumption form an upper bound

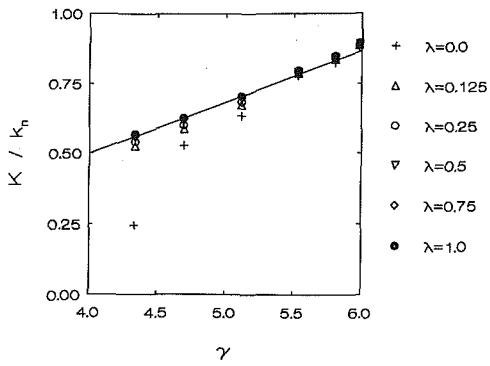


Fig. 7 Comparison between theoretical bulk modulus and bulk modulus as determined from computer simulations for various coordination numbers and stiffness ratios

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

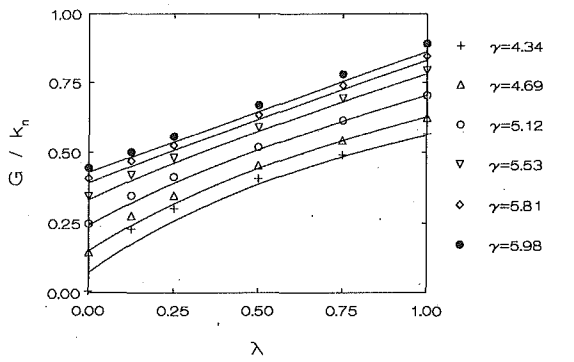


Fig. 8 Comparison between theoretical shear modulus and shear modulus as determined from computer simulations for various coordination numbers and stiffness ratios

to the elastic moduli, while the elastic moduli based on the uniform strain assumption form a lower bound to the elastic moduli. These observations are rigorously proven in Rothenburg and Kruyt (1994).

The bulk and shear modulus as determined from the computer simulations are shown in Figs. 7 and 8, together with the theoretical predictions according to (54) and (55). The agreement is fairly good, although larger deviations occur for low coordination number γ and low stiffness ratio λ.

Discussion

A micromechanical expression for the average strain tensor has been developed in the two-dimensional case, similar to an existing micromechanical expression for the average stress tensor. These relations exhibit the following duality:

$$f_i^c \leftrightarrow \Delta l_i^c$$

$$l_j^c \leftrightarrow h_j^c$$

$$\frac{1}{S}\sum_{c\in S} f_i^c l_j^c = \overline{\sigma}_{ij} \leftrightarrow \overline{\epsilon}_{ij} = \frac{1}{S}\sum_{c\in S} \Delta l_i^c h_j^c.$$

The micromechanical definition of the strain tensor has been applied to the development of (heuristic) theoretical expressions for the elastic moduli of two-dimensional, isotropic assemblies of bonded, nonrotating particles. These expressions give the moduli in terms of the coordination number γ and contact stiffnesses k_n and k_t. In contrast to Bathurst and Rothenburg (1988a, b) it has been possible to derive the dependencies of the elastic moduli K and G on coordination number γ. The developed theory has been compared with results from computer simulations; agreement is fairly good over the range of coordination numbers and stiffness ratios considered. The largest deviations occur in the case of zero tangential stiffness k_t and coordination number γ around 4. The reason for these deviations is that the system is least stable under these conditions.

The assumption of uniform strain is frequently made in continuum mechanics of heterogeneous systems (for example, Batchelor and O'Brien, 1977). This study and a related study (Rothenburg et al., 1987) show that the uniform strain assumption does not always lead to accurate prediction of the system behavior. In fact, two regimes have been distinguished here, uniform stress and uniform strain, each with its range of validity.

Future work is directed towards development of theories for the elastic moduli of bonded systems *with* particle rotations. Another subject of research is the extension of the presented micromechanical strain definition to the three-dimensional case.

References

Batchelor, G. K., and O'Brien, R. W., 1977, "Thermal or electrical conduction through a granular material," *Proceedings of the Royal Society of London,* Vol. A335, pp. 313–333.

Bathurst, R. J., and Rothenburg, L., 1988a, "Micro-mechanical aspects of isotropic granular assemblies with linear contact interactions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 17–23.

Bathurst, R. J., and Rothenburg, L., 1988b, "Note on a random isotropic granular material with negative Poisson's ratio," *International Journal of Engineering Science,* Vol. 26, pp. 373–383.

Cundall, P. A., and Strack, O. D. L., 1979, "A discrete numerical model for granular assemblies," *Géotechnique,* Vol. 29, pp. 47–65.

Cundall, P. A., Drescher, A., and Strack, O. D. L., 1982, "Numerical experiments on granular assemblies: Measurements and observations," *IUTAM Conference on Deformation and Failure of Granular Materials,* Delft, The Netherlands, pp. 355–370.

Drescher, A., and De Josselin de Jong, G., 1972, "Photoelastic verification of a mechanical model for the flow of a granular material," *Journal of the Mechanics and Physics of Solids,* Vol. 20, pp. 337–351.

Horne, M. R., 1965, "The behaviour of an assembly of rotound, rigid, cohesionless particles I and II," *Proceedings of the Royal Society of London,* Vol. A286, pp. 62–97.

Mehrabadi, M. M., Loret, B., and Nemat-Nasser, S., 1993, "Incremental constitutive relations for granular materials based on micromechanics," *Proceedings of the Royal Society of London,* Vol. A441, pp. 433–463.

Quickenden, T. I., and Tan, G. K., 1974, "Random packing in two dimensions and the structure of monolayers," *Journal of Colloid and Interface Science,* Vol. 48, pp. 382–393.

Rothenburg, L., 1980, "Micro-mechanics of idealised granular materials," Ph.D. Thesis, Department of Civil Engineering, Carleton University, Ottawa, Ontario, Canada.

Rothenburg, L., and Bathurst, R. J., 1989, "An analytical study of induced anisotropy in idealised granular materials," *Géotechnique,* Vol. 49, pp. 601–614.

Rothenburg, L., Bathurst, R. J., and Dusseault, M. B., 1989, "Micro-mechanical ideas in constitutive modelling of granular materials," *Powders and Grains,* J. Biarez and R. Gourvès, eds., Balkema, Rotterdam, The Netherlands, pp. 355–363.

Rothenburg, L., and Kruyt, N. P., 1994, "Statistical theory for the elastic moduli of two-dimensional assemblies of granular materials," in preparation.

Rothenburg, L., Matyas, E. L., and Ambrus, S. Z., 1987, "Statistical aspects of flow in a random network of channels," *Stochastic Hydrology and Hydraulics,* Vol. 1, pp. 217–240.

Rothenburg, L., and Selvadurai, A. P. S., 1981, "A micro-mechanical definition of the Cauchy stress for particulate media," *Proceedings of the International Symposium on Mechanical Behaviour of Structured Media,* A. P. S. Selvadurai, ed., Ottawa, Canada, pp. 469–486.

Satake, M., 1992, "A discrete mechanical approach to granular materials," *International Journal of Engineering Science,* Vol. 30, pp. 1525–1533.

Strack, O. D. L., and Cundall, P. A., 1978, "The distinct element method as a tool for research in granular media: Part I," Report to the National Science Foundation, NSF Grant ENG75-20711.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

C. R. Wassgren

C. E. Brennen
Mem. ASME.

M. L. Hunt
Mem. ASME.

California Institute of Technology,
Pasadena, CA 91125

Vertical Vibration of a Deep Bed of Granular Material in a Container

A deep bed of granular material (more than six layers of particles) was subjected to sinusoidal, vertical vibrations. Several phenomena were observed depending on the amplitude of excitation. These included heaping, surface waves, and arching; the transitions from one state to another involved various dynamic instabilities and bifurcations. The paper includes a description of these phenomena and the characteristic properties associated with each in addition to measurements of the transitions from one phenomena to another.

1 Introduction

A granular material flow typically consists of a particulate solid in an interstitial fluid, usually air, which is subjected to a shearing force. The response of the material is generally quite complicated and is not well understood. To better understand the behavior of these flows, several "benchmark" experiments have been investigated. These include flows in a vertical chute, down an inclined chute, between a fixed wall and moving wall (Couette flow), and in a vibrating container. These benchmark experiments are useful in that they provide a flow environment which is relatively straightforward to model when checking the validity of granular flow theories and simulations. This paper describes a detailed experimental investigation of a container of granular material subjected to vertical oscillation.

In addition to being a useful benchmark experiment, vibration of granular material is of interest for more practical reasons. In industry, vibration is commonly used as a means of handling and transporting particulate materials such as foodstuffs, coal, and pharmaceuticals. Examples include vibrating conveyor belts, hoppers, sorting tables, packing tables, drying plates, and fluidized bed reactors. Vibration of a granular material may also play an important role in natural events such as earthquakes and avalanches. Clearly, understanding how a granular material responds when subjected to vibration can provide valuable design information. More information concerning some of the important issues, questions, and applications of the knowledge of the granular state can be found in a recent review by Jaeger and Nagel (1992).

The standard granular material vibration experiment consists of a granular material such as sand or glass spheres, in a rectangular or cylindrical container mounted on a rigid base which is subjected to vertical, sinusoidal oscillations. This simple experiment exhibits some surprisingly complex behaviors that have yet to be fully explained. These include states in which the particle bed behaves as a "cloud" of particles with little or no structure, and states where the particle bed moves as a coherent mass and exhibits complex wave phenomena. A brief review of these behaviors is given in the following paragraphs.

First, it is important to review some of the variables that will influence the state of the particle bed. These include the amplitude, a, and the radian frequency of vibration, ω, the particle properties such as the particle diameter, d, and the initial particle bed depth, h_0. These variables can be combined to form two important nondimensional parameters, the dimensionless bed depth, h_0/d, which is a measure of the number of particle layers in the bed, and the dimensionless acceleration amplitude, $\Gamma = a\omega^2/g$, where g is the acceleration due to gravity. The behavior of the particle bed is strongly related to both of these dimensionless parameters.

For shallow particle beds ($h_0/d < 4$) at dimensionless acceleration amplitudes between $\Gamma = 2.5$ and $\Gamma = 6.0$, the bed exhibits three states which depend on h_0/d (see Thomas et al., 1989). In the first, or "Newtonian-I," state ($h_0/d = 0.17$ for example) particles bounce around so randomly that the vertical concentration of particles changes very little during a vibration cycle. In the second, or "Newtonian-II," state (for example at $h_0/d = 0.273$) a dense layer of particles accumulates on the vibrating surface once each vibration cycle while the remainder of the particles bounce around randomly. In the third, or "coherent-expanded" state (for example at $h_0/d = 1.7$), particles begin to move as a coherent mass which expands and contracts considerably during each vibration cycle. When h_0/d is increased above 4.0, the particle bed behavior transitions to what is called a "coherent-condensed," or deep bed state. Here particles move as a coherent mass but remain fairly compact throughout the cycle of vibration. Bachmann (1940) found that the transition from the coherent-expanded state to the coherent-condensed state occurred for large spherical particles when $h_0/d \approx 6$. In this deep bed regime, investigators have found that a number of new behaviors appear. These include the formation of heaps, surface waves, and arches.

The "heaping" phenomenon was first observed by Faraday (1831). He found that when sand is placed on a vibrating membrane, small mounds form and a slow bulk convection of particles appears. The same behavior is observed for a granular material in a container performing vertical, sinusoidal oscillations. Above a critical dimensionless acceleration amplitude of $\Gamma = 1.2$ (see, for example, Evesque and Rajchenbach, 1989), the free surface of the particle bed begins to form a mound or heap. Generally a single peak is observed to occur either to one side of the container or in the center (Fauve et al., 1989). However, if the vibration level is increased suddenly, multiple peaks sometimes appear (Fauve et al., 1989). The heap is associated with a slow particle convection pattern in which particles avalanche down the free surface of the mound, are subducted into the mass at its lowest point, and then recirculate internally back to the peak.

Although it is perhaps the most extensively studied deep-bed vibration phenomenon, the mechanism for heaping and its

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, July 18, 1994; final revision, Oct. 23, 1995. Associate Technical Editor: W. N. Sharpe, Jr.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

associated convection pattern are not well understood. Several theories explaining the behavior have been offered but none seem to be universally accepted. The reader is referred to papers by Laroche et al. (1989, 1990), Evesque and Rajchenbach (1989), Evesque (1990, 1992), Gallas et al. (1992), and Pak et al. (1995) for further details.

Another phenomenon that is observed for deep beds is the formation of surface waves. Pak and Behringer (1992) (and Behringer, 1993) observed solitary traveling waves that appear during the heaping phenomenon. The waves travel from the lowest point of the heap up the slope to the peak but do not interfere with the continuous avalanche of particles associated with the convection pattern. The waves increase in length and decrease in height as they travel up the slope and eventually disappear at the peak. Though heaping depends only on Γ, the onset of the traveling surface waves also depends on the vibration-amplitude-to-particle-diameter ratio, a/d.

Fauve et al. (1989) and Melo et al. (1994) also observed that when the acceleration level of the external oscillations is increased to the point at which heaping disappears, localized surface waves appear. For certain ranges of the excitation amplitude and frequency, standing surface waves form at half the excitation frequency. Fauve et al. suggest that the formation of these waves occurs because of a resonance between the characteristic particle avalanche time and the vibration period. This mechanism is similar to the Faraday instability (Douady and Fauve, 1988; Miles and Henderson, 1990). Melo et al. studied these waves in three-dimensional containers, as opposed to quasi-two-dimensional containers. They found several different wave patterns, such as stripes and squares, that appear after specific vibrational acceleration amplitudes. They also made measurements of the wavelengths of the standing waves.

Another phenomenon that has been observed in deep beds is known as "arching." This behavior consists of sections of the particle bed oscillating out of phase with one another. The boundaries between these regions are known as "nodes" (called "defects" by Douady et al., 1989) and exhibit very little movement. The resulting motion resembles "collapsing arches." In addition to the bulk oscillating motion of the particle bed, there is a strong convection cell pattern in which particles move down at the nodes and up at the antinodes (the regions between nodes where the peaks and troughs of the arches occur). A more complete description of the individual particle and bed motion is given by Douady et al. (1989). This phenomenon has also been observed in numerical simulations performed by Rosato and Lan (1994).

The arching behavior appears only after the motion of the particle bed undergoes a period doubling bifurcation, which occurs at $\Gamma \approx 4.5$, a value which varies with the depth of the bed (Douady et al., 1989). Prior to the bifurcation, the particle bed has a single flight time that is less than the oscillation period of the base and the bed repeats the same motion every oscillation cycle. After the bifurcation, however, the bed has two flight times: one of which is larger than the oscillation period and one which is less than the oscillation period. Thus the motion of the particle bed is repeated every two cycles instead of every cycle. Since there are two flight times, sections of the bed can be simultaneously undergoing different flight times, producing the arching motion that is observed experimentally (see Section 3.3 for more detail).

Douady et al. (1989) also found that the number of nodes observed decreases with the height of the bed (h_0). Furthermore, they found that the number of nodes is not determined by the dimensionless acceleration amplitude since different numbers of nodes can be observed for the same value of Γ. Instead there is a maximum number of nodes that can exist at a particular Γ which implies a minimum distance between the nodes. Measurements by Douady et al. indicate that the minimum distance between nodes is proportional to the bed depth, h_0, and is inversely proportional to the dimensionless acceleration amplitude, Γ. Douady et al. also report that other bifurcations occur at higher acceleration levels.

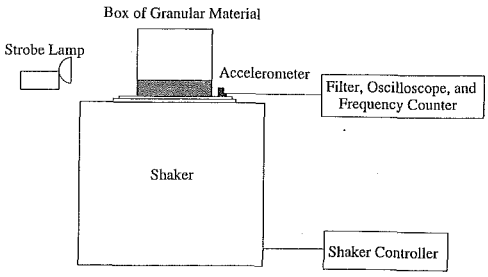


Fig. 1 A schematic of the experimental setup

The present paper presents further studies of the deep-bed phenomena of heaping, surface waves, and arching. Both qualitative and quantitative measurements are reported as well as studies of the transitions between phenomena.

2 Experimental Apparatus

Figure 1 shows a schematic of the experimental setup. The apparatus consisted of a box mounted on a Ling electromagnetic shaker (model A-175) which provided sinusoidal, vertical vibrations. The frequency of the oscillations could be controlled to within ± 0.5 Hz while the acceleration level could be controlled to within $\pm 0.1g$. The acceleration of the base of the container was monitored using a Dytran 3126A accelerometer.

Three rectangular lucite boxes were constructed to contain the granular material. The interior vertical sides of the boxes were lined with smooth window glass in order to reduce electrostatic effects. The internal, cross-sectional dimensions were 13.5 cm by 1.6 cm, 18.0 cm by 1.6 cm, and 18.0 cm by 4.8 cm. The boxes were mounted on a base of 2.5 cm thick steel.

Three different sizes of glass spheres (with a density of 2.5 g/cm^3) were used. For convenience they are referred to as particles A, B, and C and have the following properties:

Particles A: dyed (fluorescent yellow and black) glass spheres with a mean diameter of 1.28 mm (and a standard deviation of 0.04 mm) and a mean sphericity, defined as the largest diameter divided by the smallest diameter, of 1.04 (with a standard deviation of 0.03). The angle of repose of the material, θ_r, was 30.1 deg.

Particles B: 3-mm diameter glass spheres with an angle of repose of $\theta_r = 31.1$ deg.

Particles C: 3-mm diameter glass spheres with an angle of repose of $\theta_r = 26.4$ deg.

Detailed measurements of particles B and C were not made; however, the diameters of particles C were more uniform than particles B and the sphericities of C were much closer to unity.

All of the experiments were performed with air as the interstitial fluid. In the literature, uncertainty exists as to how important the interstitial fluid is in the formation of the deep-bed phenomena. Recent work by Pak et al. (1995) regarding heap formation indicates that the interstitial fluid is very important for the formation of heaps and the associated convection pattern (although they do mention that other mechanisms such as wall friction or inhomegenous forcing can still produce the observed heap and convection). They also observed that the interstitial gas effects were most important for deep beds and for small particles (glass spheres, $d < 1$ mm) which is in accord with the studies of interstitial fluid effects by Zeininger and Brennen (1985). In Pak et al.'s experiments with 1-mm diameter glass spheres and $h_0/d = 100$, the interstitial fluid effects were weak and for $d = 1.3$ mm and $h_0/d = 77$ the effects were not observed at all. Laroche et al. (1989) also found that interstitial fluid effects

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

were important for the formation of heaps for particle beds comprised of glass spheres with diameters between 0.63 and 0.8 mm but they do not mention how beds with larger particles are affected. Evesque (1990), however, reports that in experiments using 0.2, 0.4, and 1-mm diameter glass spheres the interstitial fluid effects were negligible. These experiments did not show any appreciable difference in heap shape or convection pattern for evacuated containers or for containers with mesh walls that allow air to be drawn in from the sides instead of through the particle bed. Despite the disagreement as to the whether or not fluid effects are important for the formation of the bed phenomena for small particles, the experiments performed here used particles and bed depths ($d \geq 1.28$ mm, $h_0/d < 80$) that are beyond the regime where Pak et al. regard fluid effects as being important and where Laroche et al. reported their findings. As an additional note recent discrete element simulation work (fluid effects were not included) performed by the authors have reproduced all of the mentioned bed phenomena except for the heaping behavior (Wassgren et al., 1995) further indicating that the phenomena observed here were not the result of interstitial fluid effects.

At a given frequency, f, data were obtained for acceleration levels ("g" levels), $\Gamma = a\omega^2/g$, between zero and the maximum limit of the shaker (roughly $\Gamma = 8.0$). For each experiment the acceleration amplitude and frequency were determined from the accelerometer signal. The motion of the particle mass was observed using a strobe lamp while spatial measurements were made using a grid taped to the outside of the rear wall of the box. The flight time of the particle bed (the time between successive collisions) could be measured since collisions between the particle bed and the steel base produced a narrow band of high-frequency noise in the accelerometer output. Thus, by using a high-pass filter, it was possible to measure the time between collisions of the bed with the base.

3 Observations

The transition from the coherent-expanded state to the coherent-condensed state was determined by observing when the deep-bed phenomena first appeared as Γ was increased. It was found that this transition depended on the dimensionless particle bed depth, h_0/d, and the type of material. For particles A, the critical h_0/d at which the transition occurred was between 5.5 and 8.3. For particles B, the critical h_0/d was between 5.0 and 7.5, and for particles C it was between 8.5 and 9.7.

The deep-bed phenomena observed in the present experiments included heaping, small-amplitude surface waves, arching, and large-amplitude surface waves. Two regimes of standing surface waves, defined as small and large-amplitude surface waves, were observed and were similar to the surface waves reported by Fauve et al. (1988) and Melo et al. (1994, 1995). The traveling surface waves described by Pak and Behringer (1989) were not observed in the present experiments. The sections that follow provide more detailed descriptions of the phenomena.

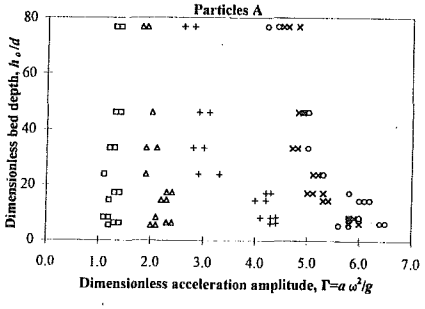


Fig. 2 Incipient values of Γ for the various phenomena as a function of h_0/d. Inception values for heaping ($\square = \Gamma_H$), small-amplitude surface waves ($\triangle = \Gamma_S$), period doubling bifurcation ($+ = \Gamma_{B1}$), return to single flight time ($\times = \Gamma^*_{B1}$), and large-amplitude surface waves ($\bigcirc = \Gamma_L$) are shown.

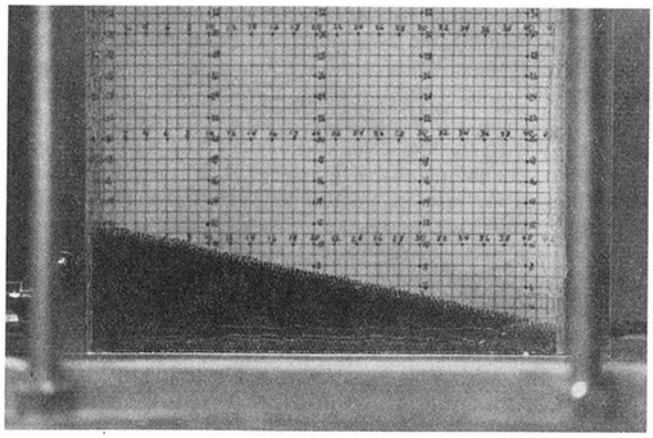


Fig. 3 A photograph of a heap for $\Gamma = 1.8$, $f = 30$ Hz, $d = 1.28$ mm, and $h_0/d = 17.5$

In addition, measurements were made of the transitions from one phenomenon to another. These transitions were found to be independent of the both the container size and the excitation frequency, f, within the $15 \rightarrow 40$ Hz range, but dependent on the dimensionless acceleration amplitude, Γ, and the number of particle layers, h_0/d. This agrees with the results of other experimenters (see, for example, Fauve et al., 1989; Evesque and Rajchenbach, 1989; Melo et al., 1995). Phenomena inception values for particles A are shown in Fig. 2. Although the value at which the transitions between phenomena occurred were different for the three types of particles, the general trends remained the same.

3.1 Heaping. A typical example of a heap is shown in Fig. 3. The inception value for heaping was $\Gamma_H \approx 1.2$ and was independent of h_0/d and particle type (Fig. 2). These results agree with observations of other experimenters; for example Evesque and Rajchenbach (1989) found that $\Gamma_H = 1.2 \pm 0.1$. Below the inception value for heaping, the surface of the particle bed remained flat with little particle motion occurring although particles occasionally fell into neighboring voids.

For $\Gamma > \Gamma_H$, the initially flat surface of the particle mass started to deform at the side boundaries and the disturbance propagated toward the center of the mass. One side of the particle mass always became thinner while the other side became thicker. The final result was a mound with the peak near a side boundary and a relatively uniform slope across most of the width of the box. Figure 4 indicates the evolution of the heap. For reasons that are obscure, the peak always formed toward the right side of the container for frequencies below 15 Hz and toward the left side for frequencies above 15 Hz. Multiple peaks and peaks in the center of the box such as described by Fauve et al. (1989) were never observed. Furthermore, when the mound was perturbed with a stirring rod the peak would always reform in its original location. Perhaps, small, unintentional lateral vibrations of the shaker caused this preferential asymmetry (see Evesque, 1992). Asymmetries in the container construction were ruled out by rotating the container 180 deg about a vertical line and observing the that the heap behavior was the same. However, when the container was rotated just 90 deg about a vertical line, the heap showed preferential peaking with a different critical frequency (note that the support structure for the shaker is not radially symmetric). Measurements of the lateral vibrations, for the container mounted in its original location, with an accelerometer indicated maximum accelerations of approximately $0.4g$ at a frequency of 30 Hz and a vertical acceleration of $1.8g$. For all other vibration parameters the

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

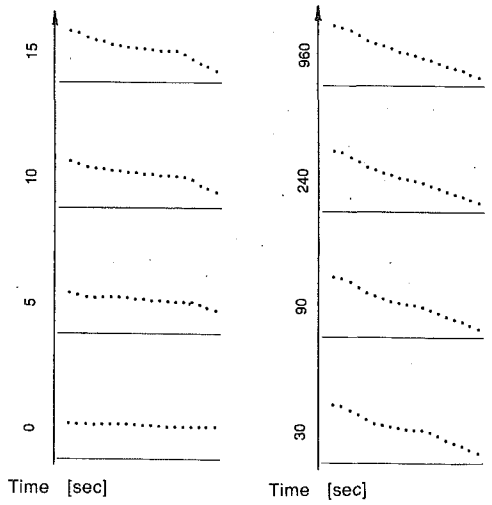


Fig. 4 The evolution of a heap over time. Data for $\Gamma = 1.4$, $f = 30$ Hz, $d = 1.28$ mm, and $h_0/d = 17.5$.

lateral acceleration was smaller than $0.4g$. The maximum final heap slope occurred at this maximum lateral acceleration. To further clarify whether or not these lateral vibrations were affecting the slope of the heap, a discrete element simulation incorporating both vertical and horizontal vibrations was performed. The resulting bed behavior showed a distinct peak to one side of the container and a convection pattern similar to what was observed experimentally.

To document one case (for particles A and $h_0/d = 17.5$), the height of the particle bed was measured at several locations along the width of the 13.5 cm by 1.6 cm box at various times during the evolution of the heap. The mean heap slope at a particular time was determined by a straight line fit through the measured heights at the prescribed locations. Figure 5 presents a typical set of experimental measurements. In all of the runs, the slope increased rapidly at first but then appeared to change much more slowly and approach some final slope. As the acceleration amplitude of the oscillations increased, the final heap slope decreased while the time required to reach the final slope increased.

A distinct convection pattern was always present during the heaping process. Particles avalanched down the slope to its lowest point and were then recirculated within the particle mass back to the peak. Although the peak of the mound was located at one end of the container, it was separated from the wall by a distance of a few particle diameters. Particles next to this boundary moved down along the wall and recirculated back to the peak. The speed of the convective motion appeared to decrease as the slope of the heap decreased implying that the speed of the convective motion decreases with increasing acceleration amplitude.

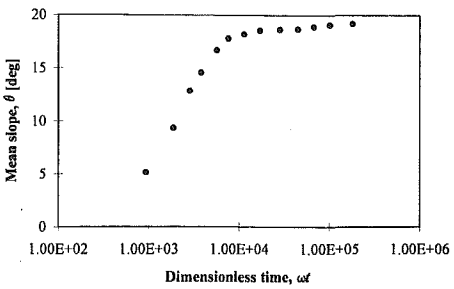


Fig. 5 The mean slope of the heap, θ, as a function of dimensionless time, ωt, at $\Gamma = 1.4$, $f = 30$ Hz, $d = 1.28$ mm, and $h_0/d = 17.5$

When the motions of the particles during each cycle were studied it was observed that during the time of flight, the particles moved as a single mass with little fluidization. They appeared to be roughly arranged in a hexagonal packing structure and had little translational motion but did rotate in place. When in flight the bottom surface of the particle mass was almost flat but with a slight curvature with the points of least clearance occurring at the walls. Laroche et al. (1989) reported that the smallest clearances occurred beneath the shallowest part of the bed whereas in the present experiments, the smallest clearances occurred beneath the shallowest and deepest parts, both of which were located next to the walls.

3.2 Small-Amplitude Surface Waves. Small-amplitude surface waves began to appear on the slope of the heap when Γ was increased above another critical value denoted by Γ_s; $\Gamma_s \approx 2.2$ but depended on both h_0/d and particle type (see Fig. 2). As Γ was increased further, the heap and its convection pattern disappeared and the surface waves became more pronounced. Figure 6 is a typical photograph of small-amplitude surface waves.

The small-amplitude waves usually appeared as two-dimensional standing waves with crests parallel to the short dimension of the box although more complicated wave patterns sometimes occurred (for example, sloshing kink waves and cnoidal waves, Guthart, 1992; Miles and Henderson, 1990). Sometimes different portions of the surface would have different standing wave patterns. For example, one end might have waves with crests parallel to the short dimension of the box while the other end might have waves with the crests parallel to the long dimension. When the crests were all parallel to the short dimension, the number of waves varied between n and $n + \frac{1}{2}$ where n is an integer. The walls always coincided with antinodes and the wave motion repeated every two vibration cycles. These characteristics are similar to those found with the parametrically excited Faraday instability waves in liquids (see Miles and Henderson, 1990). Figure 7 shows the wave motion and the trajectories of particles within the waves.

The peaks and troughs of the waves were rounded for small vibration amplitudes, a, but the peaks became sharper and the troughs broader as the amplitude increased. In fact as the amplitude increased, the peaks became sharp enough so that particles were ejected from the peaks during the wave motion and would land in the adjacent troughs. Only the top 1–2 layers of particles remained fluidized throughout the wave motion. Below those upper layers, the next 2–4 layers of particles seemed to remain in contact with neighboring particles, but did show movement relative to one another due to the shearing associated with the wave motion. The remaining particles below these upper 3–6

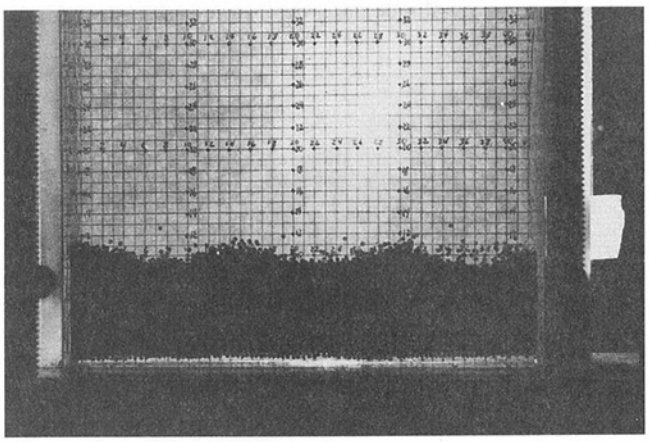


Fig. 6 A photograph of small-amplitude surface waves at $\Gamma = 3.3$, $f = 20$ Hz, $d = 1.28$ mm, and $h_0/d = 17.5$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

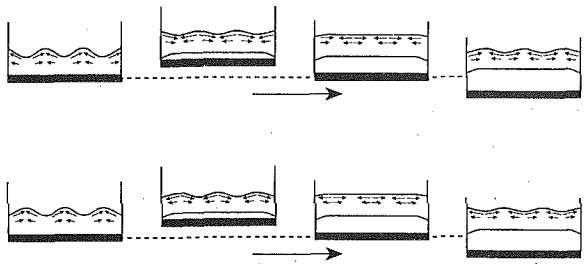


Fig. 7 A schematic showing the motion of the particles for small-amplitude surface waves during two oscillation cycles

layers appeared to remain in a packed state. Note that these observations were made at the vessel walls rather than in the bulk of the material. As Γ increased, the fluidized particles in the top few layers became increasingly more active. Although the wave motion was still evident, the fluidized particles "smeared" the waves.

In flight, the bottom surface of the particle mass had a shape similar to that which occurred during heaping; it remained flat with a slight decrease in the clearance near the walls. In the shallow particle beds the waves seemed to disappear near the value of Γ at which the flight time of the particle bed bifurcated (Γ_{B1}). However, in the deeper particle beds they persisted beyond Γ_{B1}.

Measurements were made of the wave amplitude, η (defined as half of the maximum elevation difference between the peaks and troughs), and the wavelength, λ, for particles A in the 13.5 cm by 1.6 cm box with $h_0/d = 17.5$. An appropriate dimensionless parameter commonly used in fluid dynamics when investigating surface waves is the Froude number which can be interpreted as the ratio of a characteristic velocity of a flow to the surface wave speed. In Fig. 8 the Froude number based on the wave amplitude, η,

$$Fr = \frac{a\omega}{\sqrt{g\eta}} \tag{1}$$

is plotted against Γ. This figure shows that the wave amplitude correlates well with this Froude number. The Froude number increases with Γ starting at $Fr = 1.0$ and remains greater than 1 throughout the range of Γ investigated.

Wavelength measurements were attempted by determining the distance between neighboring wave peaks. This method, however, was not very accurate since the waves peaks were rounded and the precise location of the peak was difficult to determine. Therefore only a few qualitative comments will be made. It appeared that the wave amplitude and wavelength both increased as h_0/d increased but then reached a critical value of h_0/d where the amplitude and wavelength remained constant.

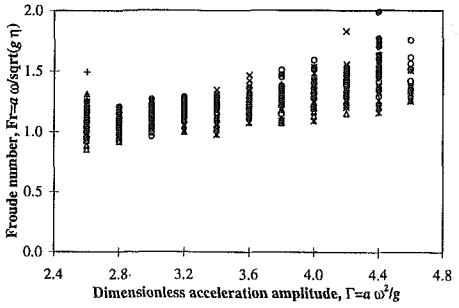


Fig. 8 Froude number, Fr = $a\omega/\sqrt{g\eta}$, based on the wave amplitude, η, plotted against the dimensionless acceleration amplitude, Γ, for small-amplitude surface waves (d = 1.28 mm and h_0/d = 17.5). Frequency key: + = 15 Hz, ● = 20 Hz, × = 25 Hz, ○ = 30 Hz, △ = 35 Hz, * = 40 Hz.

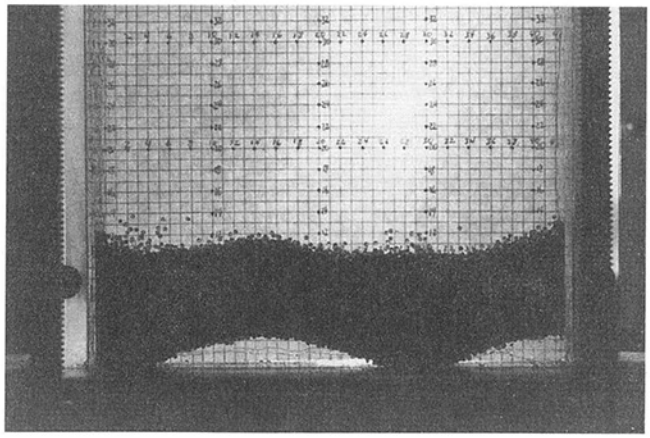


Fig. 9 A photograph of 3-node arching with a single flight time (Γ = 8.5, f = 30 Hz, d = 1.28 mm, and h_0/d = 17.5)

This wavelength observation is similar to observations made by Melo et al. (1994) who found that the wavelength increased as h_0/d increased but remained constant for values of $h_0/d > 7$. Other observations indicated that the wavelength increased with increasing particle diameter and was relatively independent of the container width as long as the container width is greater than a wavelength. Both of these observation are consistent with the observations made by Melo et al. (1990).

3.3 Post Bifurcation — Arching. When the acceleration level reached a particular bifurcation value, Γ_{B1}, the motion of the particle bed underwent a period doubling bifurcation with respect to the imposed frequency. Instead of repeating the same motion every oscillation cycle, the motion repeated every two cycles. In this regime, sections of the particle bed could oscillate out-of-phase producing the behavior known as arching (see Fig. 9). The boundaries between the sections, are termed nodes and oscillated with the smallest amplitude during the motion of the mass. The term antinode is used for the regions between nodes which had the largest amplitudes.

The formation of arches can be visualized by modeling the bed as two completely inelastic particles separated by a fixed horizontal length and vibrating vertically on a sinusoidally oscillating table. Consider first a single inelastic particle. Between collisions with the table, the vertical position, y, of the particle follows a ballistic trajectory,

$$\eta = -\frac{1}{2}(\phi - \phi_{n-1})^2 + \dot{\eta}_{n-1}(\phi - \phi_{n-1}) + \eta_{n-1}, \tag{2}$$

where the subscript $(n-1)$ refers to values immediately following the previous collision and the equation is nondimensionalized as follows: $\eta = \omega^2 y/g$, $\dot{\eta} = \omega\dot{y}/g$, and $\phi = \omega t$. Collisions between the floor and the particle occur when

$$-\frac{1}{2}(\phi_n - \phi_{n-1})^2 + \dot{\eta}_{n-1}(\phi_n - \phi_{n-1}) + \eta_{n-1} - \Gamma \sin(\phi_n) = 0 \tag{3}$$

where ϕ_n indicates the dimensionless time of the next collision.

Since the particle is completely inelastic it assumes the velocity of the floor following a collision. If the acceleration of the floor is greater than $-g$ then the particle remains on the floor until the floor acceleration falls below $-g$. By analyzing these equations of motion, one finds that a period doubling bifurcation occurs when $\Gamma_{B1} \approx 3.7$ (Holmes, 1984; Mehta and Luck, 1990). The subscript $B1$ refers to the first period-doubling bifurcation. Other period-doubling bifurcations occur at larger Γ but only the first will be considered here.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

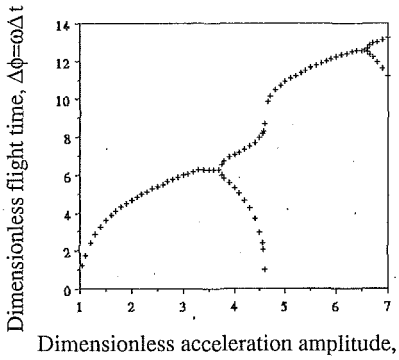


Fig. 10 The dimensionless flight time, $\Delta\phi = \omega\Delta t$ of a completely inelastic particle vibrating on a sinusoidally oscillating table plotted as a function of dimensionless acceleration amplitude, $\Gamma = a\omega^2/g$. The first period doubling bifurcation occurs at $\Gamma_{B1} \approx 3.7$ and the return to a single flight time is at $\Gamma^*_{B1} \approx 4.6$. A second period doubling bifurcation occurs at $\Gamma_{B2} \approx 6.6$.

Another critical value of Γ occurs when $\Gamma^*_{B1} \approx 4.6$. The significance of both Γ_{B1} and Γ^*_{B1} becomes apparent by considering two particles on a vibrating table. For $\Gamma < \Gamma_{B1}$, both particles collide with the floor every oscillation cycle after a flight time of Δt which is less than or equal to the period of the forcing oscillations, $T = 2\pi/\omega$ (see Fig. 10). If one of the particles is perturbed, it will return to the same motion it had prior to the perturbation. Thus a line connecting the two particles remains horizontal. For $\Gamma_{B1} < \Gamma < \Gamma^*_{B1}$ the particles undergo a period-doubling bifurcation and their motion is repeated every two oscillation cycles. Here the particles have two flight times namely Δt_1, which is larger than T and Δt_2 which is less than T (the sum $\Delta t_1 + \Delta t_2$ is less than $2T$). Both particles can either be undergoing the same flight time concurrently (in-phase) or can be performing different flight times (out-of-phase). When in-phase, a line connecting the two remains horizontal. On the other hand the out-of-phase motion produces a line that tilts from one side to the other depending on the location of the particles. Thus two different types of motion are exhibited for the same value of Γ. This same behavior is observed for an entire particle bed. The in-phase motion is known as 0-node arching since no nodes are present and the out-of-phase motion is called n-node arching where n is the number of nodes present.

As Γ increases, Δt_1 increases and Δt_2 decreases until the next critical value $\Gamma = \Gamma^*_{B1}$ at which the two flights merge resulting in a single flight time occurring every two cycles. Under these conditions ($\Gamma^*_{B1} < \Gamma < \Gamma_{B2}$ where the second period doubling bifurcation occurs at $\Gamma_{B2} \approx 6.6$) only one flight time, $\Delta t_1 < 2T$, remains but the motion of the particle bed still repeats every two oscillation cycles. Again the particle motions can be either in-phase or out-of-phase.

As an interesting side note, a similar bifurcation also appears for much shallower particle beds (for example $h_0/d \approx 0.5$) but produces different results. In Brennen et al. (1992) the authors measured the height of a moveable lid containing a shallow bed of particles as a function of Γ. The height increased suddenly at a critical Γ. By examining the dynamics of an *elastic* particle bouncing on a sinusoidally oscillating table (as opposed to the completely inelastic particle just examined), one finds a bifurcation at a particular value of Γ. In that case the bifurcation caused a "jump" in the maximum height which the particles achieved. Thus, bifurcations occur for both shallow and deep particle beds and can be explained by similar particle bed dynamics by using different effective coefficients of restitution for the collisions between the particles and the base.

Returning to the present observations it was observed that for all three granular materials, the values of Γ_{B1} and Γ^*_{B1} decreased as h_0/d increased and then reached a constant value

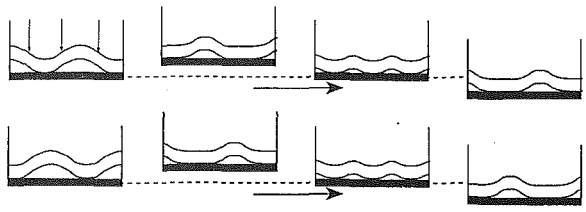


Fig. 11 A schematic showing the particle bed motion for 3-node arching. Arrows in the first frame indicate the location of the nodes.

(see Fig. 2). For particles A, Γ_{B1} and Γ^*_{B1} decreased from values of 4.2 and 5.8 for $h_0/d = 5.5$ and reached constant values of 2.9 and 4.5 for $h_0/d \approx 30$. The difference, $\Gamma_{B1} - \Gamma^*_{B1}$, remained relatively constant at about 1.6. Particles B and C followed the same trends. The relationship between Γ_{B1} and h_0/d is the reverse of that reported by Douady et al. (1989). They found that Γ_{B1} increased with increasing h_0/d and suggested that the cause was due to the damping of the interstitial fluid. This discrepancy might be due to the difference in particle sizes used in the experiments. In the experiments by Douady et al., glass spheres with diameters between 0.63 and 0.8 mm were used (salt was also used), within the regime where interstitial fluid effects are important (Pak et al., 1995). The present experiments, however, used particles with diameters of 1.28 mm and 3 mm, for which the interstitial fluid effects are negligible.

When $\Gamma_{B1} < \Gamma < \Gamma^*_{B1}$, either zero, one, or, occasionally, two nodes were present and particles within the mass moved in slow convective cell motions. As Γ was increased, the phase difference between the neighboring sections, the amplitude of the arches, and the particle motion within the mass all increased. When $\Gamma^*_{B1} < \Gamma < \Gamma_{B2}$, neighboring sections of the particle bed oscillated 180 deg out-of-phase. Figure 11 shows the particle bed motion for arching.

At a given Γ the number of nodes was not unique. However, a critical Γ exists below which a particular number of nodes never occurred. This result is equivalent to the finding by Douady et al. (1989), which indicated the existence of a minimum allowable distance between nodes, l. In their experiments using glass spheres with diameters between 0.63 and 0.80 mm, they found $h_0/l \approx 0.16(\Gamma - 4.2)$. In the present experiments similar measurements were made for particles A and are presented in Fig. 12. Clearly h_0/l increases with Γ.

The nodes were positioned symmetrically within the box when three or more nodes were present and the vertical walls always coincided with antinodes. For one or two nodes the arch could be quite asymmetric. In fact, using a stirring rod to perturb the particles, one could change the location of the nodes within the container.

At a given Γ, the number of nodes could be decreased by perturbing the particle mass. As Γ increased, the motion of the

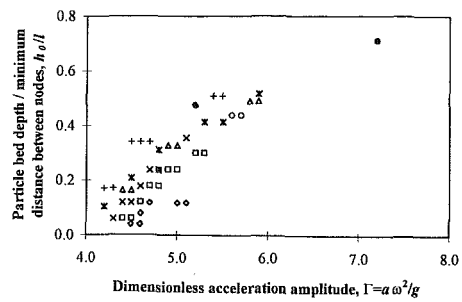


Fig. 12 The particle bed depth, h_0, divided by the minimum distance between nodes, l, as a function of dimensionless acceleration amplitude, Γ, using particles A with $f = 30$ Hz and $d = 1.28$ mm. Legend: ◇ = $h_0/d = 5.5$, □ = $h_0/d = 6.3$, × = $h_0/d = 8.3$, ∗ = $h_0/d = 14.5$, △ = $h_0/d = 17.2$, + = $h_0/d = 23.8$, ● = $h_0/d = 33.3$, and ○ = $h_0/d = 46.1$.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

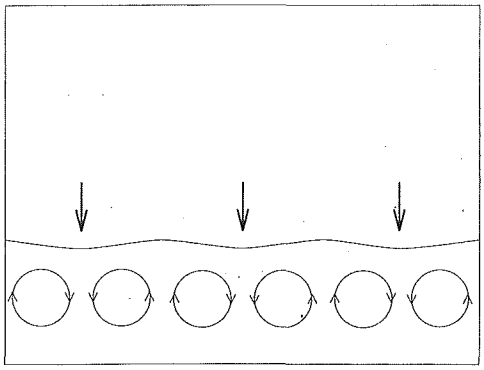


Fig. 13 Particle motion for 3-node arching averaged over several oscillation cycles. The orientation of the convection cells is indicated by the arrows on the circles and the thick arrows indicate the nodes.

particle mass sometimes became erratic, transitioning between 0-node arching and n-node arching (n is an integer greater than zero). Occasionally the motion settled into arching with a larger number of nodes than before the erratic motion. For shallower beds, the arching behavior with a large number of nodes was more stable than with a small number of nodes. However, as the particle bed became deeper 0-node arching became the more stable motion.

Particles within the mass moved in circular convection cells when averaged over several oscillation cycles, moving down at the nodes and up at the antinodes (see Fig. 13). The number of convection cells was twice the number of nodes. The particle speed increased with increasing Γ and with an increase in the number of nodes but decreased as the frequency increased. The descriptions by Douady et al. (1989) of the particle motions are consistent with the present observations.

For the special case of 0-node arching, one large convection cell was present similar to the convection pattern observed during heaping. A shallow slope was usually seen beneath a fluidized surface. The undersurface of the particle bed was similar to that observed during heaping; it remained flat with a slight curvature at the walls. Similar experimental observations have been made by Ehrichs et al. (1995); others have observed these features in discrete element simulations (Taguchi, 1992; Gallas et al., 1992; Wassgren et al., 1995). It should be noted that the style of shaking differs somewhat in these various studies. For example, Ehrichs et al. use discrete taps to shake the container instead of continuous, sinusoidal vibration. Moreover, a number of the simulations have fixed walls and a sinusoidally oscillating floor. Convection is observed in all of these studies despite these differences.

3.4 Large-Amplitude Surface Waves. When Γ increased above another critical value, Γ_L, large amplitude surface waves appeared superimposed on the arching behavior (see Fig. 14). As shown in Fig. 2, Γ_L occurred very close to the value of $\Gamma^*_{\beta 1}$. As with $\Gamma^*_{\beta 1}$, Γ_L decreased with increasing h_0/d and eventually reached a constant value near $h_0/d \approx 30$.

Large-amplitude waves only formed in regions of the particle bed where the top surface was essentially flat. Thus, waves were seen only during zero, one, and two node arching. For the most part the behavior of the large-amplitude waves was similar to that of the small-amplitude waves (see Section 3.2); however, several differences were apparent. First, the peaks of the large amplitude waves were much sharper than the rounded peaks of the small amplitude waves. Also, the large amplitude waves always formed as standing waves with crests parallel to the short dimension of the container and were large enough that particles were always ejected from the peaks of the waves, landing in the neighboring troughs. Lastly, instead of repeating the same wave motion every two oscillation cycles, the motion was repeated every four cycles.

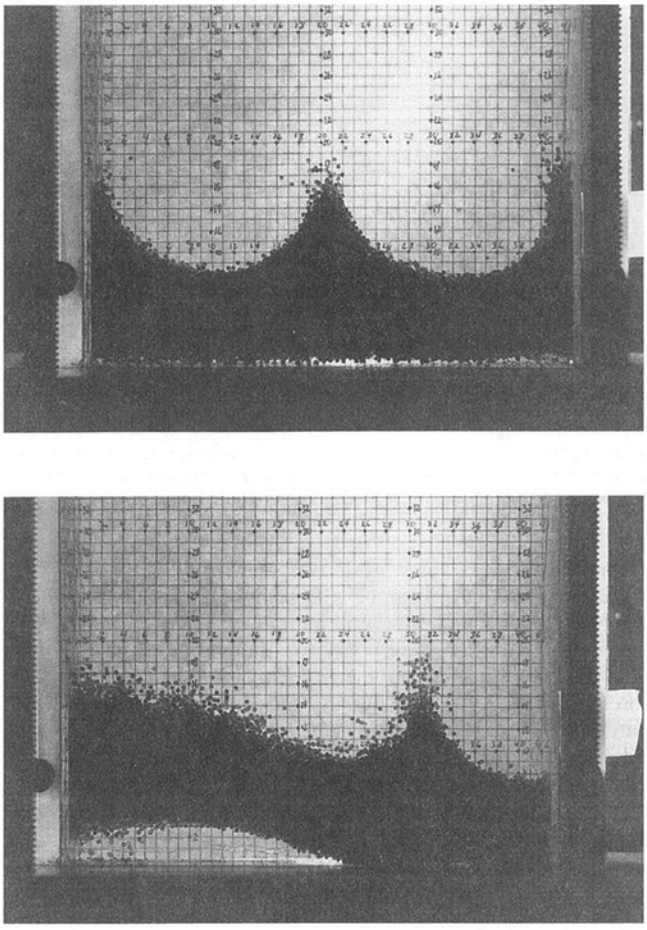


Fig. 14 Two photographs of large-amplitude surface waves. The top photo shows waves superimposed on 0-node arching and the bottom photo shows waves superimposed on 1-node arching for $\Gamma = 6.2$, $f = 20$ Hz, $d = 1.28$ mm, and $h_0/d = 17.5$.

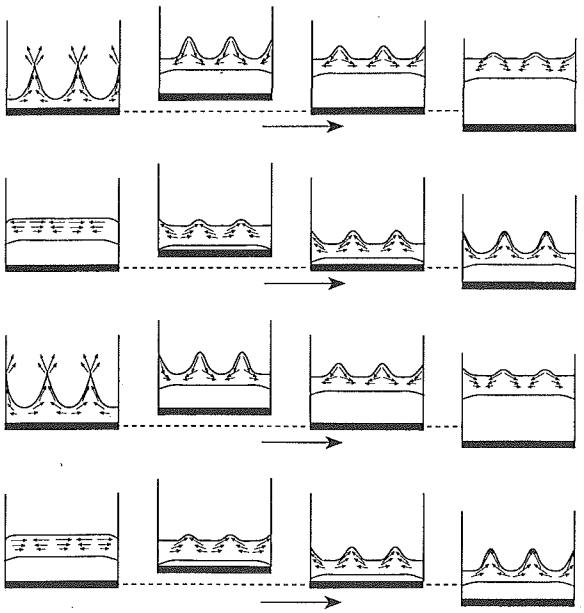


Fig. 15 A schematic showing the motion of the waves and particles for large-amplitude surface waves and 0-node arching

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

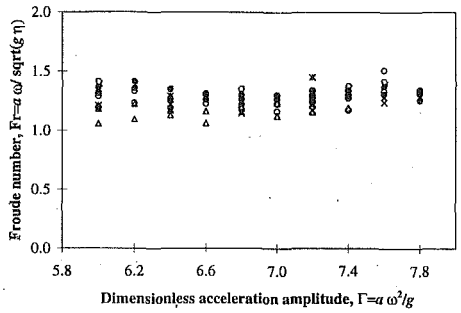


Fig. 16 Froude number, Fr = $a\omega/\sqrt{g\eta}$, based on the wave amplitude, η, plotted against the dimensionless acceleration amplitude, Γ, for large-amplitude surface waves (d = 1.28 mm and h_0/d = 17.5). Frequency key as in Fig. 8.

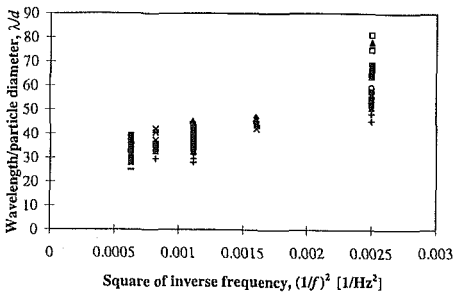


Fig. 17 Wavelength, λ, plotted against the square of the inverse frequency, $(1/f)^2$, for various dimensionless acceleration amplitudes, Γ, for large-amplitude surface waves (d = 1.28 mm and h_0/d = 17.5). Dimensionless acceleration amplitude (Γ) key: + = 6.0, ● = 6.2, × = 6.4, ○ = 6.6, △ = 6.8, * = 7.0, − = 7.2, ■ = 7.4, ♦ = 7.6, and □ = 7.8.

The particle motions within the mass were similar to those observed with small-amplitude waves. Particles at the surface were fluidized while within the mass they remained in a packed state but moved relative to one another due to the shearing of the wave motion which was much greater than for small-amplitude waves. The number of waves present was given by $n/2$ and the lateral boundaries always coincided with antinodes. Figure 15 shows a schematic of the wave motion with 0-node arching.

Measurements were made of the wave amplitude, η, and the wavelength, λ, for particles A in the 13.5 cm by 1.6 cm box with $h_0/d = 17.5$. Again the amplitude data correlates well with the Froude number of Eq. (1); Fig. 16 shows that Fr is close to constant at a value of about 1.2; note that this is the same as for the small-amplitude waves.

Wavelength measurements were made by measuring the distance between peaks and are more accurate than the small-amplitude data since the peaks are well defined and concave. The wavelengths for various acceleration amplitudes are plotted against the square of the inverse frequency in Fig. 17 in order to compare with the measurements made by Melo et al. (1994). Clearly the wavelength data correlates well with the square of the inverse frequency, as was the case with the data of Melo et al.

As with the small-amplitude waves (see Section 3.2), both the amplitude and wavelength appear to increase with increasing h_0/d but then saturate at a critical value of h_0/d. Additionally, as Γ is increased, the waves begin to "tilt" alternately from side to side with neighboring peaks either pointing toward or away from one another.

4 Conclusions

When a vessel containing a deep bed of granular material is vibrated, a wide range of different flow phenomena can occur depending on the level of the imposed vibration. These phenomena are heaping, surface waves, and arching. Each of these involves complex two and three-dimensional particle motions. Many of the details of the particle motions remain to be understood. This paper provides some quantitative documentation of these complex dynamical phenomena and, in some cases, a tentative framework for an understanding of the bifurcations that occur in the dynamics.

5 Acknowledgments

The authors wish to acknowledge the support of the National Science Foundation who supported this work under NSF grant CTS 9300665.

References

Bachmann, D., 1940, *Verfahrenstechnik Z.V.D.I. Beiheft,* No. 2, p. 43, (cited by Gutman, 1976a).

Behringer, R., 1993, "The dynamics of flowing sand," *Nonlinear Science Today,* Vol. 3, No. 3, pp. 1–15.

Brennen, C., Ghosh, S., and Wassgren, C., 1993, "Vertical Oscillation of a Bed of Granular Material," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 156–161.

Douady, S., and Fauve, S., 1988, "Pattern selection in Faraday instability," *Europhys. Lett.,* Vol. 6, No. 3, pp. 221–226.

Douady, S., Fauve, S., and Laroche, C., 1989, "Subharmonic instabilities and defects in a granular layer under vertical vibrations," *Europhys. Lett.,* Vol. 8, No. 7, pp. 621–627.

Ehrichs, E., Jaeger, H., Karczmar, G., Knight, J., and Kuperman, V., 1995, "Granular convection observed by magnetic resonance imaging," *Science,* Vol. 267, No. 5204, pp. 1632–1634.

Evesque, P., 1990, "Comment on: Convective flow of granular masses under vertical vibrations," *J. Phys. France,* Vol. 51, No. 8, pp. 697–699.

Evesque, P., 1992, "Shaking dry powders and grains," *Contemp. Phys.,* Vol. 33, No. 4, pp. 245–261.

Evesque, P., and Rajchenbach, J., 1989, "Instability in a sand heap," *Phys. Rev. Lett.,* Vol. 62, No. 1, pp. 44–46.

Faraday, M., 1831, "On a peculiar class of acoustical figures and on certain forms assumed by groups of particles upon vibrating elastic surfaces," *Phil. Trans. R. Soc. London,* Vol. 52, pp. 299–340.

Fauve, S., Douady, S., and Laroche, C., 1989, "Collective behaviors of granular masses under vertical vibration," *J. Phys. France,* Vol. 50, No. 3, pp. 187–191.

Gallas, J., Herrmann, H., and Sokolowski, S., 1992, "Convection cells in vibrating granular media," *Phys. Rev. Lett.,* Vol. 69, No. 9, pp. 1371–1374.

Guthart, G., 1992, "On the existence and stability of standing solitary waves in Faraday resonance," Ph.D. Thesis, California Institute of Technology, Pasadena, CA.

Holmes, P., 1982, "The dynamics of repeated impacts with a sinusoidally vibrating table," *J. of Sound and Vibration,* Vol. 84, No. 2, pp. 173–189.

Jaeger, H., and Nagel, S., 1992, "Physics of the granular state," *Science,* Vol. 255, No. 5051, pp. 1523–1531.

Laroche, C., Douady, S., and Fauve, S., 1990, "Convective flow of granular masses under vertical vibrations," *J. Phys. France,* Vol. 50, No. 7, pp. 699–706.

Laroche, C., Douady, S., and Fauve, S., 1990, "Reply to Comment on: Convective flow of granular masses under vertical vibrations," *J. Phys. France,* Vol. 51, No. 8, p. 700.

Mehta, A., and Luck, J., 1990, "Novel temporal behavior of a nonlinear dynamical system: The completely inelastic bouncing ball," *Phys. Rev. Lett.,* Vol. 65, No. 4, pp. 393–396.

Melo, F., Umbanhowar, P., and Swinney, H., 1994, "Transition to parametric wave patterns in a vertically oscillated granular layer," *Phys. Rev. Lett.,* Vol. 72, No. 1, pp. 172–175.

Melo, F., Umbanhowar, P., and Swinney, H., 1995, "Hexagons, kinks, and disorder in oscillated granular layers," *Phys. Rev. Lett.,* Vol. 75, No. 21, pp. 3838–3841.

Miles, J., and Henderson, D., 1990, "Parametrically forced surface waves," *Annual Rev. Fluid Mech.,* Vol. 22, pp. 143–165.

Pak, H., and Behringer, R., 1993, "Surface waves in vertically vibrated granular materials," *Phys. Rev. Lett.,* Vol. 71, No. 12, pp. 1832–1835.

Pak, H., Van Doorn, E., and Behringer, R., 1995, "Effects of ambient gases on granular materials under vertical vibration," *Phys. Rev. Lett.,* Vol. 74, No. 23, pp. 4643–4646.

Rosato, A., and Lan, Y., 1994, "Granular dynamics modeling of vibration-induced convection of rough inelastic spheres," *Proceedings of the First International Particle Technology Forum,* Denver, CO.

Thomas, B., Mason, M., Liu, Y., and Squires, A., 1989, "Identifying states in shallow vibrated beds," *Powder Tech.,* Vol. 57, pp. 267–280.

Wassgren, C. R., Brennen, C. E., and Hunt, M. L., 1995, "Discrete element simulations of a deep bed of granular material subjected to vertical, sinusoidal vibrations," manuscript in preparation.

Zeininger, G., and Brennen, C. E., 1985, "Interstitial fluid effects in hopper flows of granular material," ASME *Cavitation and Multiphase Flow Forum,* Vol. 23, pp. 132–136.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Modeling of Cyclic Ratchetting Plasticity, Part I: Development of Constitutive Relations

Y. Jiang[1]
Research Assistant.

H. Sehitoglu
Professor.

Department of Mechanical and Industrial Engineering,
University of Illinois at Urbana-Champaign,
1206 W. Green Street,
Urbana, IL 61801

The existing plasticity models recognize that ratchetting direction strongly depends on the loading path, the stress amplitude, and the mean stresses, but their predictions deviate from experiments for a number of materials. We propose an Armstrong-Frederick type hardening rule utilizing the concept of a limiting surface for the backstresses. The model predicts long-term ratchetting rate decay as well as constant ratchetting rate for both proportional and nonproportional loadings. To represent the transient behavior, the model encompasses a memory surface in the deviatoric stress space which recalls the maximum stress level of the prior loading history. The coefficients in the hardening rule, varying as a function of the accumulated plastic strain, serve to represent the cyclic hardening or softening. The stress level effect on ratchetting and non-Masing behavior are realized with the size of the introduced memory surface. Simulations with the model checked favorably with nonproportional multiaxial experiments which are outlined in Part 2 of this paper.

1 Introduction

Plasticity models or constitutive equations are mathematical relations describing the stress-strain response of a material subjected to external loading. For engineering purposes, when phenomenological constitutive equations are developed, simplifications in material description are adopted. For example, the material is considered to be homogenous and initially isotropic. In reality, at the micro-level, metals are comprised of many grains with different orientations, associated phases, and dislocation structures which are dynamically evolving under stressing. In addition, in this study, we will consider the regimes of loading where the response is time-independent and isothermal. The response of materials at elevated temperatures ($>0.35\ T_m$ where T_m is the melting temperature) and under thermomechanical loading may require further experiments and modified constitutive descriptions (Slavik and Sehitoglu, 1986).

This investigation deals with constitutive models in the stress space. In the presence of external stress and multiaxial stress states, the behavior of the material can be rather complex, especially when loading paths are nonproportional. The limitations of the previous models have been demonstrated in earlier work (Jiang and Sehitoglu, 1994a, 1994b), particularly for ratchetting type loadings. In the case of ratchetting, the hysteresis loops do not close at the conclusion of the cycle, i.e., if the plastic strain in the loading portion of the cycle exceeds that under the unloading portion, one gets "forward" ratchetting and the opposite case is referred to as "backward" ratchetting. The accumulation of plastic strain occurs over repeated contacts of rail and wheel in railroads and contributes to material flow and wheel/rail cracking.

Experimental ratchetting behaviors of 1070 steel (0.7 percent carbon steel used in railroads) under changing stress paths and multiaxial loads have been investigated by the authors (Jiang and Sehitoglu, 1994a, 1994b) and the reasons for different ratchetting trends have been explained. We showed that the hypothesis that *ratchetting occurs in the direction of nonzero mean stress* does not always hold true even for proportional loading. For example, in a two-step loading test, when the mean stresses in the second loading step are decreased slightly but keep the same signs as in the first step, the ratchetting may occur in the opposite directions to the mean stresses. Ratchetting rate may decrease, reach a minimum and then increase again during constant amplitude loading. The material memory, which contributes to these experimental ratchetting phenomena, dissipates with increasing number of cycles. A comprehensive model should have the capabilities to reproduce all these trends. Yet, most existing plasticity models are unable to capture the different ratchetting behaviors. It was shown that the models often miss even the ratchetting direction (Jiang and Sehitoglu, 1994a, 1994b). Ultimately, the constitutive models will be used in inelastic analyses of components where mean stresses and nonproportional loading are present and ratchetting and fatigue result in failure. A small difference in plastic strain rates in one cycle between model and experiment could translate to large deviations over many cycles. Therefore, it is imperative that these models predict accurate inelastic strains under various loading conditions.

After a critical review of the framework of plasticity models, including the concept of limiting surfaces for the backstresses, we will introduce a new plasticity model in this paper. The new model will be used to predict cyclic ratchetting results where accumulation of plastic strain over many cycles develops.

1.1 Yield Criteria and Plastic Modulus Function. During elastic-plastic deformation, the Von Mises yield function is employed,

$$f = (\mathbf{S} - \boldsymbol{\alpha}):(\mathbf{S} - \boldsymbol{\alpha}) - 2k^2 = 0, \qquad (1)$$

where $\mathbf{S}$ is the stress tensor in the deviatoric space, $\boldsymbol{\alpha}$ is the deviatoric backstress (center of the yield surface), and k is the yield stress in simple shear. A bold letter denotes a second-order Cartesian tensor. A colon between two second-order tensors denotes their scalar product. In Fig. 1, a geometric illustration of the yield surface is depicted with kinematic translation in the deviatoric stress space. The plastic strain increment is

[1] Current address: Department of Mechanical Engineering, University of Nevada, Reno, NV 89557.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, June 20, 1994; final revision, Apr. 28, 1995. Associate Technical Editor: W. N. Sharpe, Jr.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

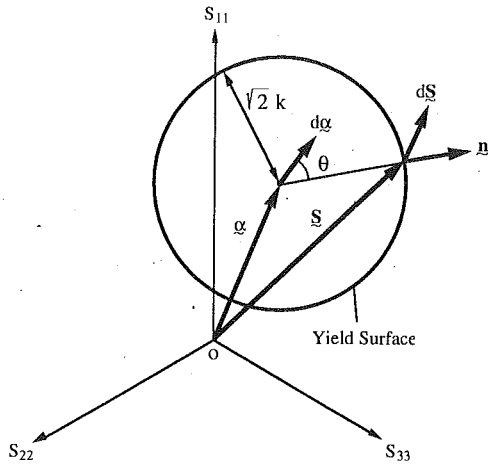


Fig. 1 Yield surface and yield surface translation

collinear with the exterior normal to the yield surface at the present stress state. This normality flow rule can be mathematically expressed as

$$d\boldsymbol{\epsilon}^p = \frac{1}{h}\langle d\mathbf{S} : \mathbf{n}\rangle \mathbf{n}, \tag{2}$$

where $\langle\ \rangle$ denotes the MacCauley bracket (i.e., $\langle x\rangle = (x + |x|)/2$), $\mathbf{n}$ is the unit exterior normal to the yield surface at the loading point, and h is a scalar function often called the plastic modulus function.

The consistency condition, which requires that the stress state lies on the yield surface, should be satisfied during elastic-plastic deformation. Using the consistency condition, we can obtain the plastic modulus function,

$$h = \frac{\mathbf{n} : d\boldsymbol{\alpha}}{dp} + \sqrt{2}\,\frac{dk}{dp}, \tag{3}$$

where $dp = \sqrt{d\boldsymbol{\epsilon}^p : d\boldsymbol{\epsilon}^p}$ is generally called the equivalent plastic strain increment. Equation (3) determines the plastic modulus function, h, when the evolution of the backstress and the yield stress have been specified.

Since the plasticity theories will be discussed in a stress-based space, proportional loading is defined as the loading corresponding to a fixed radial line in the deviatoric stress space throughout the entire loading history. Furthermore, a loading under which a virgin material initially isotropic will undergo no ratchetting and/or mean stress relaxation is referred to as "balanced" loading. Consequently, an "unbalanced" loading is prescribed as the loading under which a virgin material initially isotropic will undergo strain ratchetting and/or stress relaxation.

1.2 A Review of Hardening Rules in View of Ratchetting Data. A hardening rule specifies changes in the yield condition as a result of loading of the material in the elastoplastic regime. When the plastic modulus function is a constant, the linear hardening rule predicts no ratchetting for proportional loading and no steady state ratchetting for nonproportional loading. When the plastic modulus function is expressed as a function of the second deviatoric stress invariant (Drucker and Palgen, 1981), a constant ratchetting rate is predicted for both proportional and nonproportional loading. The ratchetting rate predicted is generally much larger than the experimental results (Hassan et al., 1992). These hardening rules do not correlate with experimental data for nonproportional cyclic loading (Lamba and Sidebottom, 1978; McDowell, 1983). Further studies (McDowell, 1985c; Jiang and Sehitoglu, 1994a) demonstrated that the experimental translation direction of the yield surface is different from the exterior normal to the yield surface at the stress point for nonproportional loading.

Multiple surface models (Garud, 1981; Iwan, 1967; Mroz, 1967) do not predict ratchetting for any proportional loading. However, these models predict ratchetting for general nonproportional loading (Garud, 1991; Jiang, 1993) but the predicted ratchetting rates are often much larger than the experimental observations. McDowell (1992) drew attention to the fact that the original forms of two-surface plasticity theory cannot fundamentally correlate cyclic strain accumulation under proportional and nonproportional loading over a variety of amplitudes and mean stresses.

Armstrong and Frederick (1966) developed a nonlinear kinematic hardening relation, introducing a recovery term associated with a strain memory effect,

$$d\boldsymbol{\alpha} = a_a d\boldsymbol{\epsilon}^p - c_a \boldsymbol{\alpha} dp, \tag{4}$$

where a_a is a constant, c_a is a scalar function of the plastic strain path, and dp was defined earlier. With the introduction of a nonlinear term, this model renders constant strain rate ratchetting for constant amplitude stress-controlled loading. Generally this model leads to large overestimations of ratchetting (Chaboche, 1989b; Chaboche and Nouailhas, 1989a; Clement and Guionnet, 1985; Inoue et al., 1989; Jiang, 1993; Ohno and Wang, 1993b). Bower (1989) modified the A-F (Armstrong-Frederick) rule by adding one more internal state variable and his model is able to predict ratchetting rate decay only for a limited number of cycles, after which it predicts ratchetting arrest (Jiang, 1993; Jiang and Sehitoglu, 1994a).

Nonlinear hardening rules based on the A-F relation have been expressed in the form of a series expansion of the backstress. It was postulated (Chaboche et al., 1979) that the total backstress is composed of additive parts,

$$\boldsymbol{\alpha} = \sum_{i=1}^{M} \boldsymbol{\alpha}^{(i)}, \tag{5}$$

where $\boldsymbol{\alpha}$ is the total backstress, $\boldsymbol{\alpha}^{(i)}$ is a part of the total backstress, $i = 1 \sim M$, and M is the number of backstress parts considered. Each backstress, $\boldsymbol{\alpha}^{(i)}$, follows an A-F type relation,

$$d\boldsymbol{\alpha}^{(i)} = c^{(i)}(r^{(i)} d\boldsymbol{\epsilon}^p - W^{(i)}\boldsymbol{\alpha}^{(i)} dp) \quad (i = 1, 2, \ldots, M), \tag{6}$$

where $c^{(i)}$ and $r^{(i)}$ are scalar functions associated with the ith part of backstress $\boldsymbol{\alpha}^{(i)}$, and $W^{(i)}$ is a function of stress state. Chaboche et al.'s initial model (Chaboche et al., 1979) can be represented by $W^{(i)} = 1$.

Under the asymmetric uniaxial loading this model predicts a transient ratchetting response consistent with experiments, but the duration of the transient response is short-lived. Subsequently the model predicts constant ratchetting rate (Jiang, 1993; Jiang and Sehitoglu, 1994a, 1994b).

Chaboche (1991) introduced a threshold term which allows dynamic recovery only when a backstress reaches a critical value. The introduction of the threshold produces less ratchetting than the model without the threshold. Following these works, Ohno and Wang (1991, 1993a) proposed a different threshold concept,

$$W^{(i)} = \left(\frac{|\boldsymbol{\alpha}^{(i)}|}{r^{(i)}}\right)^{\chi^{(i)}} \langle \mathbf{n} : \mathbf{L}^{(i)}\rangle \quad (i = 1, 2, \ldots, M), \tag{7}$$

where $\chi^{(i)}$ $(i = 1, 2, \ldots, M)$ are material constants, and

$$\mathbf{L}^{(i)} = \frac{\boldsymbol{\alpha}^{(i)}}{\|\boldsymbol{\alpha}^{(i)}\|} \quad (i = 1, 2, \ldots, M), \tag{8}$$

and

$$\|\boldsymbol{\alpha}^{(i)}\| = \sqrt{\boldsymbol{\alpha}^{(i)} : \boldsymbol{\alpha}^{(i)}} \quad (i = 1, 2, \ldots, M). \tag{9}$$

When the magnitude of a backstress $\boldsymbol{\alpha}^{(i)}$ is less than a thresh-

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

old value $r^{(i)}$, $W^{(i)} \approx 0$, hence the strain hardening is approximately linear. The nonlinear term in Eq. (6) is fully recovered only when $\|\boldsymbol{\alpha}^{(i)}\|$ reaches a threshold level.

From an early evaluation (Jiang and Sehitoglu, 1994a, 1994b), it was found that among many existing models, the Ohno-Wang model is the only one that can properly predict the experimentally observed ratchetting rate decay over many loading cycles. With appropriate selection of the material constants, the Ohno-Wang model can provide ratchetting predictions in close agreement with the experimental results (Jiang, 1993; McDowell, 1992). A major shortcoming of the Ohno-Wang model is its inability to predict constant ratchetting rate for nonproportional loading (Jiang and Sehitoglu, 1994a).

The development of a new plasticity model was inspired from the anomalous results of multiple step ratchetting experiments (Jiang, 1993; Jiang and Sehitoglu, 1994a, 1994b) and some success achieved with the Ohno-Wang model. The proposed model is aimed at overcoming the shortcomings of the Ohno-Wang model. Before we introduce the new model, we discuss the structure of the hardening rules and introduce the concept of limiting surface for a backstress. In Part 2 of the paper, the simulations with the proposed model will be compared to experimental results.

2 A General Class of Hardening Rules

A series expansion of backstresses, as was expressed in Eq. (5), is used in the current formulation and, it is assumed that each backstress rate takes the following form:

$$d\boldsymbol{\alpha}^{(i)} = f_1^{(i)}(\mathbf{n} - f_2^{(i)}\mathbf{L}^{(i)})dp \quad (i = 1, 2, \ldots, M), \quad (10)$$

where $\mathbf{n}$, dp, and $\mathbf{L}^{(i)}$ remain the same as defined earlier. The quantities $f_1^{(i)}$ and $f_2^{(i)}$ in Eq. (10) are scalar functions of the backstresses, which satisfy the following conditions:

Condition 1 $f_1^{(i)}$ is single-valued and non-negative, and,
Condition 2 $f_2^{(i)} \geq \mathbf{n}:\mathbf{L}^{(i)}$ when $|\boldsymbol{\alpha}^{(i)}| = r^{(i)}(r^{(i)} > 0)$.

From Eq. (10) it can be derived that

$$d\boldsymbol{\alpha}^{(i)}:\mathbf{L}^{(i)} = f_1^{(i)}(\mathbf{n}:\mathbf{L}^{(i)} - f_2^{(i)}\mathbf{L}^{(i)}:\mathbf{L}^{(i)})dp$$
$$= f_1^{(i)}(\mathbf{n}:\mathbf{L}^{(i)} - f_2^{(i)})dp. \quad (11)$$

According to Condition 1, $f_1^{(i)}$ is greater than or equal to zero. The equivalent plastic strain increment, dp, is also non-negative. Therefore from Eq. (11), the sign of the quantity $d\boldsymbol{\alpha}^{(i)}:\mathbf{L}^{(i)}$ is the same as the sign of the quantity $(\mathbf{n}:\mathbf{L}^{(i)} - f_2^{(i)})$. Knowing that $f_2^{(i)} \geq \mathbf{n}:\mathbf{L}^{(i)}$ when $|\boldsymbol{\alpha}^{(i)}| = r^{(i)}$ (Condition 2), examination of Eq. (11) leads to the following inequality:

$$d\boldsymbol{\alpha}^{(i)}:\mathbf{L}^{(i)} \leq 0 \quad \text{when} \quad |\boldsymbol{\alpha}^{(i)}| = r^{(i)}. \quad (12)$$

Since the backstresses, $\boldsymbol{\alpha}^{(i)}$, are initially zero, this previous inequality implies that whenever $\boldsymbol{\alpha}^{(i)}$ reaches a surface with a radius $r^{(i)}$, $d\boldsymbol{\alpha}^{(i)}$ makes an obtuse or right angle to $\boldsymbol{\alpha}^{(i)}$. As a result of this characteristic the backstress $\boldsymbol{\alpha}^{(i)}$ will never traverse the surface which has a radius of $r^{(i)}$ and is centered at the origin. This surface will be referred to as a limiting surface for a backstress $\boldsymbol{\alpha}^{(i)}$ in subsequent discussions. It should be noted that the limiting surface concept introduced here differs conceptionally from the limiting surface (also called bounding surface) in the two surface models.

The framework characterized by Eq. (10) with Condition 1 and Condition 2 represents a class of nonlinear hardening rules where the total backstress is divided into several parts and each part of the total backstress has a limiting surface. All of the A-F type hardening rules, including those of Chaboche et al. (1979) and Ohno-Wang (1993a, 1993b), are specific cases of this framework. The basic A-F rule, Eq. (4), is the case for which $M = 1$, $f_1^{(1)} = a_a$, $f_2^{(1)} = c_a/a_a\|\boldsymbol{\alpha}\|$. For this model, there is a limiting surface for the total backstress $\boldsymbol{\alpha}$ with a radius of $r^{(1)} = a_a/c_a$. The Chaboche hardening rule without a threshold is represented by $f_1^{(i)} = c^{(i)}r^{(i)}$ and $f_2^{(i)} = \|\boldsymbol{\alpha}^{(i)}\|/r^{(i)}$ (refer to Eq. (6) with $W^{(i)} = 1$). The Ohno-Wang model is the case where $f_1^{(i)} = c^{(i)}r^{(i)}$ and $f_2^{(i)} = (\|\boldsymbol{\alpha}^{(i)}\|/r^{(i)})^{\chi^{(i)}+1}\langle \mathbf{n} : \mathbf{L}^{(i)}\rangle$. With this generalization, the existing A-F type hardening rules can be explored with a limiting surface concept. In other words, a prevalent characteristic of all the A-F type hardening rule is the current interpretation of the limiting surface for the backstress part. The difference between the A-F type hardening rules lies in the distinctive specification of the magnitude and direction of a backstress within a limiting surface. Such an interpretation allows the material constants for these models to be determined via a standard procedure, which will be a focal point of Part 2 of this paper. More importantly, the limiting surface concept provides a convenient framework to construct refined hardening algorithms.

4 New Plasticity Model

4.1 Hardening Rule. Assume that the total backstress $\boldsymbol{\alpha}$ is divided into M parts (Eq. (5)). It is now proposed that each backstress $\boldsymbol{\alpha}^{(i)}$ takes on the following specific form:

$$d\boldsymbol{\alpha}^{(i)} = c^{(i)}r^{(i)}\left[\mathbf{n} - \left(\frac{|\boldsymbol{\alpha}^{(i)}|}{r^{(i)}}\right)^{\chi^{(i)}+1}\mathbf{L}^{(i)}\right]dp$$

$$(i = 1, 2, \ldots, M). \quad (13)$$

In general, $c^{(i)}$, $r^{(i)}$, and $\chi^{(i)}$ are three sets of non-negative and single-valued scalar functions. A geometric interpretation of the new hardening rule is presented in Fig. 2. The increment of backstress is shown for two points, namely P' and P. The line T in Fig. 2(b) denotes the tangent at point P. The limiting surface for the ith backstress part $\boldsymbol{\alpha}^{(i)}$ has a radius of $r^{(i)}$. If the exponent $\chi^{(i)} \gg 0$, the nonlinear term in Eq. (13) is approximately zero when the backstress $\boldsymbol{\alpha}^{(i)}$ is within the limiting surface ($\|\boldsymbol{\alpha}^{(i)}\| < r^{(i)}$). As a result, the increment of this backstress part, $d\boldsymbol{\alpha}^{(i)}$, follows the normal direction, $\mathbf{n}$, to the yield surface (Fig. 2(a)) for both proportional and nonproportional loading. The nonlinear term in Eq. (13) is fully recovered when the backstress is on the limiting surface ($\|\boldsymbol{\alpha}^{(i)}\| = r^{(i)}$) (Fig. 2($b$)).

Comparing Eq. (13) with the Ohno-Wang hardening rule (Eqs. (6) and (7)), we find that the difference between the two hardening rules is that the Ohno-Wang rule includes the MacCauley bracket term $\langle \mathbf{L}^{(i)}:\mathbf{n}\rangle$. The implication of this term is that when a backstress part $\boldsymbol{\alpha}^{(i)}$ is on the limiting surface, the increment of this backstress part, $d\boldsymbol{\alpha}^{(i)}$, is in the direction tangential to the limiting surface ($\mathbf{L}^{(i)}:\mathbf{n} \geq 0$) or follows the normal direction to the yield surface ($\mathbf{L}^{(i)}:\mathbf{n} < 0$). When $\boldsymbol{\alpha}^{(i)}$ is within the limiting surface, translation of the backstress part is in a direction normal to the yield surface. In contrast to the Ohno-Wang model, in the current hardening rule, $d\boldsymbol{\alpha}^{(i)}$ always makes an obtuse angle with $\boldsymbol{\alpha}^{(i)}$ for nonproportional loading when $\boldsymbol{\alpha}^{(i)}$ is on the limiting surface (refer to Fig. 2). The new hardening rule differs little from the Ohno-Wang model for the deformation modeling of proportional loading, especially when $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) are large numbers. Nonproportional loading is where the effects of this term become important. It should be noted that when $\chi^{(i)} = 0$ ($i = 1, 2, \ldots, M$) the new hardening rule delivers the initial Chaboche model (Chaboche et al., 1979) represented by Eq. (6) with $W^{(i)} = 1$. Since the Chaboche model has the ability to predict constant ratchetting rate for both proportional and nonproportional loading, the new hardening rule overcomes the Ohno-Wang model's inability to predict constant ratchetting rate for nonproportional loading. The presence of nonzero $\chi^{(i)}$ terms enables the new hardening rule to model materials that display ratchetting rate decay.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

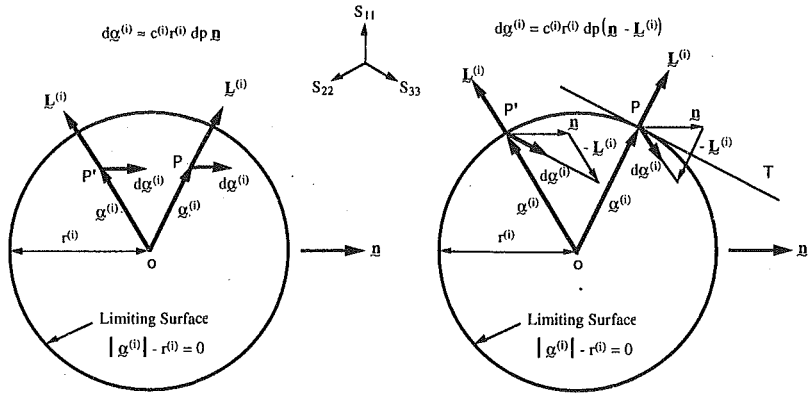


Fig. 2 Geometric interpretation of the new hardening relation; (*a*) within limiting surface, (*b*) on limiting surface

The resulting plastic modulus function corresponding to the new hardening rule is

$$h = \sum_{i=1}^{M} c^{(i)} r^{(i)} \left(1 - \left(\frac{\|\boldsymbol{\alpha}^{(i)}\|}{r^{(i)}} \right)^{\chi^{(i)}+1} \mathbf{L}^{(i)} : \mathbf{n} \right) + \sqrt{2} \frac{dk}{dp} \quad (14)$$

where the first term represents the changes in plastic modulus with backstress and the second term is the contribution of plastic modulus function due to isotropic hardening.

4.2 Memory Surface. In this study, a memory surface is introduced in the deviatoric stress space to improve the prediction of transient behavior. Transient changes in stress-strain response could occur when loading magnitudes are changed in the experiment. These changes could bring about ratchetting rate decay or acceleration. A scalar function g is used to represent this surface,

$$g = \|\boldsymbol{\alpha}\| - R_M \leq 0, \quad (15)$$

where $\|\boldsymbol{\alpha}\|$ is the magnitude of the total backstress $\boldsymbol{\alpha}$ as defined by $\|\boldsymbol{\alpha}\| = \sqrt{\boldsymbol{\alpha} : \boldsymbol{\alpha}}$. The evolution for the variable R_M, the radius of the memory surface, is

$$dR_M = H(g)\langle \mathbf{L} : d\boldsymbol{\alpha} \rangle - c_M \left\langle 1 - \frac{|\boldsymbol{\alpha}|}{R_M} \right\rangle dp, \quad (16)$$

where the tensorial quantity $\mathbf{L}$ is defined in terms of the total backstress as follows:

$$\mathbf{L} = \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}. \quad (17)$$

H in Eq. (16) represents the Heaviside step function (i.e., $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$). The initial value for R_M is zero. In numerical analysis, it is often practical to assume R_M to be a small number. A schematic illustration of the proposed memory surface is presented in Fig. 3. When the total backstress $\boldsymbol{\alpha}$ is on the memory surface and is moving outward (Fig. 3(*a*)), the memory surface expands so that the condition $g = 0$ is satisfied. When the backstress is moving within the memory surface, the memory surface may contract (if $c_M > 0$) with subsequent plastic deformation (Fig. 3(*b*)). This ability of the surface to contract mirrors experimental results where the memory of prior events decays with additional cycling. The cyclically stabilized memory surface will have a radius which is the maximum $\|\boldsymbol{\alpha}\|$ of the $\boldsymbol{\alpha}$ locus for the loading history under which the material stabilized. This geometrically implies that the memory surface and the $\boldsymbol{\alpha}$ locus inscribe at one point (Fig. 3(*c*)). The current value of R_M more strongly reflects the recent prior loading history rather than the entire history.

The introduction of new internal variables that memorize the prior maximum plastic strain range was first proposed by Chaboche et al. (1979). The concept uses a memory surface in the plastic strain space, which Ohno (1982) and Ohno and Kachi (1986) called the cyclic nonhardening range. A memory surface was also introduced in the stress space to memorize the maximum backstress of the prior history (Bruhns and Pape, 1989; Chaboche, 1989a; Mroz, 1981). The difference between the existing memory surface concepts lies in the choice of plastic strain space and stress space, the "stress/strain range" or "maximum stress/strain" to measure the memory effect, and whether there is a recovery term. With the ability to contract, the current memory surface is formalized in the stress space and the size of the memory surface characterizes the "maximum stress." The experimental observations under multiple step loading (Tanaka et al., 1985a) and non-Masing behavior (Abdel-Raouf et al., 1977; Bayerlein et al., 1987; Mughrabi, 1978; Winter, 1974) reveal that the memory of a previous higher loading amplitude fades with the subsequent number of loading cycles, suggesting that the ability for a memory surface to contract is necessary for the description of the transient behavior in many materials.

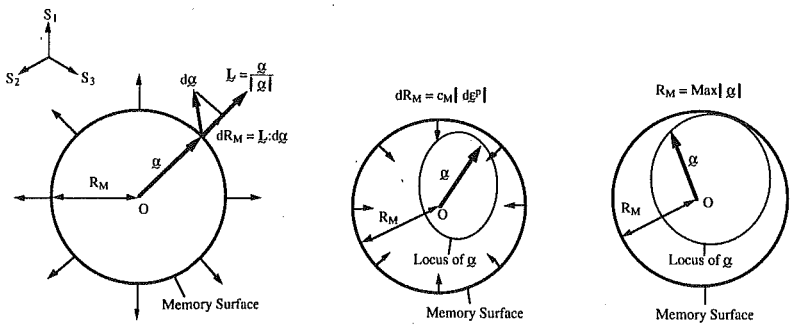


Fig. 3 Schematic illustration of memory surface; (*a*) expansion of memory surface, (*b*) contraction of memory surface, and (*c*) stabilized memory surface

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

4.3 Cyclic Hardening/Softening. Progressive elevation or decrease in flow strength with cycles results in cyclic hardening or softening, respectively. Cyclic hardening can be considered with $c^{(i)}$ in the hardening relation (Eq. (13)) being a function of the accumulated plastic strain,

$$c^{(i)} = c_0^{(i)}(1 + a_1^{(i)}e^{-b_1^{(i)}p} + a_2^{(i)}e^{-b_2^{(i)}p}) \quad (i = 1, 2, \ldots, M), \quad (18)$$

where $c_0^{(i)}$, $a_1^{(i)}$, $a_2^{(i)}$, $b_1^{(i)}$, and $b_2^{(i)}$ are material constants. When the material displays monotonic hardening or monotonic softening, only two terms in Eq. (18) are needed. When the material behavior is mixed hardening/softening, all three terms are necessary.

In a modification of the Ohno-Wang hardening rule, McDowell (1992) suggested that the cyclic hardening be considered through the coefficients $r^{(i)}$ $(i = 1, 2, \ldots, M)$ as functions of the accumulated plastic strain. In this case, there is a possibility for a backstress $\boldsymbol{\alpha}^{(i)}$ to be outside of the limiting surface, which violates Condition 2 and may lead to qualitative changes in the translation of the backstress part. To illustrate this possibility, consider a uniaxial tension-compression test. Assume that a material is undergoing elastic-plastic deformation and a backstress $\boldsymbol{\alpha}^{(i)}$ has been saturated (i.e., on a limiting surface with radius $r^{(i)}$). According to either the Ohno-Wang or the new hardening rule, $\boldsymbol{\alpha}^{(i)}$ should stay on the limiting surface as long as the uniaxial loading is maintained in the same direction. Furthermore, consider that the limiting surface is allowed to contract with plastic deformation when simulating cyclic softening. Contraction of a limiting surface when the backstress $\boldsymbol{\alpha}^{(i)}$ has been saturated could result in $\boldsymbol{\alpha}^{(i)}$ being outside of the limiting surface, if additional conditions are not enforced. In order to avoid unnecessary complexity, the use of $c^{(i)}$ (Eq. 18) when modeling this behavior eliminates a possible inconsistency.

4.4 Nonproportionality and Stress Level Effects on Ratchetting. We note that the evolution of backstress should depend on nonproportionality of loading. Moreover, the ratchetting rates may explicitly depend on the level of stress with higher degree of dynamic recovery at higher stress levels. If stress level and nonproportionality have effects on ratchetting, modification of $\chi^{(i)}$ seems appropriate. The quantity $(\mathbf{n}:\mathbf{L}^{(i)})$ can be incorporated into the formulation of $\chi^{(i)}$ to consider the nonproportionality effect on ratchetting in the following relation:

$$\chi^{(i)} = \chi_0^{(i)}(2 - \mathbf{n} : \mathbf{L}^{(i)}) \quad (i = 1, 2, \ldots, M), \quad (19)$$

where $\chi_0^{(i)}$ $(i = 1, 2, \ldots, M)$ are constants for a given stress level. The exponent $\chi_0^{(i)}$ can be further related to the size of the memory surface, R_M, to consider the stress level effect on ratchetting. The following relationship is used to model the stress level effect:

$$\chi_0^{(i)} = Q^{(i)}(1 + a_\chi e^{b_\chi R_M}) \quad (i = 1, 2, \ldots, M), \quad (20)$$

where $Q^{(i)}$ $(i = 1, 2, \ldots, M)$, a_χ, and b_χ are constants. Combining Eq. (19) with (20) results in

$$\chi^{(i)} = Q^{(i)}(2 - \mathbf{n} : \mathbf{L}^{(i)})(1 + a_\chi e^{b_\chi R_M}). \quad (21)$$

Because $(\mathbf{n} : \mathbf{L}^{(i)})$ can be interpreted as a measure of nonproportionality, the term $(2 - \mathbf{n} : \mathbf{L}^{(i)})$ takes into account the nonproportionality effect on ratchetting. It is shown in Part 2 that the larger the exponents $\chi^{(i)}$ $(i = 1, 2, \ldots, M)$ the faster the model predicts ratchetting rate decay. Equation (21) implies that the exponents $\chi^{(i)}$ $(i = 1, 2, \ldots, M)$ are larger under nonproportional loading than under proportional loading. The term $(1 + a_\chi e^{b_\chi R_M})$ in Eq. (21) accounts for the stress level effect. For higher stress levels, the memory surface is larger. As a result, $\chi^{(i)}$ is larger, and the predicted ratchetting rate will be smaller than that when $\chi^{(i)}$ takes a constant. Therefore, the predicted ratchetting rate will be smaller than when $\chi^{(i)}$ is a constant. This is consistent with the experimental observations on 1070 steel (Jiang, 1993) and those reported by Chaboche and Nouailhas (1989a,b).

4.5 Non-Masing Behavior. Non-Masing behavior refers to the nonsimilitude of the hysteresis loop shapes observed for cyclic loading experiments under different loading amplitudes. Jhansale (1975) suggested that the non-Masing behavior be modeled by considering changes in the yield stress. When refining a fatigue damage parameter Ellyin (1985) also used a similar concept to incorporate non-Masing behavior. The non-Masing results on various materials (Abdel-Raouf et al., 1977; Jhansale, 1975; Winter, 1974) indicate that the change in yield stress is a reasonable choice to model this phenomenon. Because the non-Masing behavior is related to the recent loading history, the size of the memory surface can be used to model this behavior. The following relationship is suggested:

$$k = k_1(1 + a_k e^{c_k R_M}), \quad (22)$$

where k_1, a_k, and c_k are constants. The simulations of non-Masing behavior using Eq. (22) in conjunction with the new hardening rule was confirmed to be able to capture the phenomenon both qualitatively and quantitatively (Jiang, 1993).

5 Discussion of the Model and Concluding Remarks

An attempt has been made towards a better description of cyclic ratchetting. Based on the fundamental framework of incremental plasticity theory, including yield surface, yield surface translation, and normality flow rule, a concept of the limiting surface for a backstress is generalized from the Armstrong-Frederick type hardening rules and used in the development of a new model.

The parameter $(\mathbf{n} : \mathbf{L}^{(i)})$ was used to characterize the nonproportionality effect on ratchetting. For proportional loading $(\mathbf{n} : \mathbf{L}^{(i)})$ is unity and for nonproportional loading it is less than unity. The size of the memory surface was used to consider the stress level effect on both ratchetting and non-Masing behavior. Transient cyclic behaviors are then incorporated into the hardening rule. Cyclic hardening is considered through the coefficients $c^{(i)}$ $(i = 1, 2, \ldots, M)$ in the hardening rule as power functions of the accumulated plastic strain. The coefficients $r^{(i)}$ $(i = 1, 2, \ldots, M)$ represent the radii of the limiting surfaces, which are assumed constant irrespective of the loading condition. The other coefficients, $\chi^{(i)}$ $(i = 1, 2, \ldots, M)$, can be further functions of the stress state to consider the transient cyclic behavior and are related to the size of the memory surface to capture the stress level effect. Because $\chi^{(i)}$ $(i = 1, 2, \ldots, M)$ have influence only for the unbalanced loading, $c^{(i)}$ $(i = 1, 2, \ldots, M)$ considers all the cyclic transient material properties observed in balanced loading.

The non-Masing behavior is considered through the yield stress as function of the memory surface size. The major difference between the model developed and the one proposed by Ohno and Wang lies in the ability of the new model to predict constant ratchetting rate for nonproportional loading. The model predictions are compared to experimental results in Part 2 of this paper, and its capabilities are demonstrated.

The model does not consider the cross-hardening effect observed in some materials. Following Benallal and Marquis (1987), a reasonable variable which can be related to a memory parameter for describing the cross-hardening effect could be $(\mathbf{L}:\mathbf{n})$. A parameter memorizing $\|\mathbf{L}:\mathbf{n}\|$, which fades with loading history, may be constructed in an analogous fashion to that forwarded for the memory surface described by Eqs. (15) and (16). Additional coefficients in the $c^{(i)}$ $(i = 1, 2, \ldots, M)$ format could be added for this memory parameter, which could then consider the cross-hardening effect.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Acknowledgments

Professor Sehitoglu acknowledges the financial support from Department of Energy, Basic Energy Sciences, DOE DEFG02-93ER14393. The continued support of Association of American Railroads, Technical Center, Chicago, IL, and the encouragement of Dr. Dan Stone, AAR, is also acknowledged.

References

Abdel-Raouf, H., Topper, T. H., and Plumtree, A., 1977, "Cyclic Plasticity and Masing Behaviour in Metals and Alloys," *Fracture 1977,* ICF4, Vol. 2, Waterloo, Canada, pp. 1207–1215.

Armstrong, P. J., and Frederick, C. O., 1966, "A Mathematical Representation of the Multiaxial Bauschinger Effect," Report RD/B/N 731, Central Electricity Generating Board.

Bayerlein, M., Christ, H.-J., and Mughrabi, H., 1987, "A Critical Evaluation of the Incremental Step Test," *Low Cycle Fatigue and Elasto-Plastic Behaviour of Materials,* Rie, ed., Elsevier, Amsterdam, pp. 149–154.

Benallal, A., and Marquis, D., 1987, "Constitutive Equations for Nonproportional Elastio-Viscoplasticity," ASME *Journal of Engineering Materials and Technology,* Vol. 109, pp. 326–336.

Bower, A. F., 1989, "Cyclic Hardening Properties of Hard-Drawn Copper and Rail Steel," *Journal of Mechanics and Physics of Solids,* Vol. 37, pp. 455–470.

Bruhns, O. T., and Pape, A., 1989, "A Three Surface Model in Nonproportional Cyclic Plasticity," *Proceedings of International Conference on Constitutive Laws for Engineering Materials,* Chonqing, China, Fan, and Murakami, eds., Vol. 2, Pergamon, pp. 703–708.

Chaboche, J. L., Dang Van, K., and Cordier, G., 1979, "Modelization of the Strain Memory Effect on the Cyclic Hardening of 316 Stainless Steel," *Structural Mechanics in Reactor Technology,* Transactions of the Fifth International Conference on Structural Mechanics in Reactor Technology, Div. L, Berlin, L11/3.

Chaboche, J. L., 1989a, "A New Kinematic Hardening Rule with Discrete Memory Surfaces," *Rech. Aerosp.,* No. 4, pp. 49–69.

Chaboche, J. L., 1989b, "A New Constitutive Framework to Describe Limited Ratchetting Effects," *Plasticity '89,* Pergamon, pp. 211–214.

Chaboche, J. L., 1991, "On Some Modifications of Kinematic Hardening to Improve the Description of Ratchetting Effects," *International Journal of Plasticity,* Vol. 7, pp. 661–678.

Chaboche, J. L., and Nouailhas, D., 1989a, "Constitutive Modeling of Ratchetting Effects—Part I: Experimental Facts and Properties of the Classical Models," ASME *Journal of Engineering Materials and Technology,* Vol. 111, pp. 384–392.

Chaboche, J. L., and Nouailhas, D., 1989b, "Constitutive Modeling of Ratchetting Effects—Part II: Possibilities of Some Additional Kinematic Rules," ASME *Journal of Engineering Materials and Technology,* Vol. 111, pp. 409–416.

Chaboche, J. L., Nouailhas, D., Pacou, D., and Paulmier, P., 1991, "Modeling of the Cyclic Response and Ratchetting Effects on Inconel-718 Alloy," *European Journal of Mechanics,* A/Solids, Vol. 10, pp. 101–121.

Clement, G., and Guionnet, C., 1985, "Analysis of Biaxial Experiments on 316 Steels at Elevated Temperature," *Trans. SMIRT 8,* Vol. 1, Brussels, Belgium.

Drucker, D. C., and Palgen, L., 1981, "On Stress-Strain Relations Suitable for Cyclic and Other Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 479–485.

Ellyin, F., 1985, "Effect of Tensile-Mean-Strain on Plastic Strain Energy and Cyclic Response," ASME *Journal of Engineering Materials and Technology,* Vol. 107, pp. 119–125.

Garud, Y. S., 1981, "Multiaxial Fatigue of Metals," Ph.D Dissertation, Mechanical Engineering, Stanford University.

Garud, Y. S., 1991, "Notes on Cyclic Dependent Ratchetting under Multiaxial Loads Including Bauschinger Effect and Non-Linear Strain Hardening," *Transactions of the 11th International Conference on Structural Mechanics in Reactor Technology,* Tokyo, Japan, Shibata, ed., Vol. L, pp. 511–518.

Hashiguchi, K., 1993a, "Fundamental Requirements and Formulation of Elastoplastic Constitutive Equations with Tangential Plasticity," *International Journal of Plasticity,* Vol. 9, pp. 525–549.

Hashiguchi, K., 1993b, "Mechanical Requirements and Structures of Cyclic Plasticity Models," *International Journal of Plasticity,* Vol. 9, pp. 721–748.

Hassan, T., and Kyriakides, S., 1992, "Ratchetting in Cyclic Plasticity, Part I: Uniaxial Behavior," *International Journal of Plasticity,* Vol. 8, pp. 91–116.

Hassan, T., Corona, E., and Kyriakides, S., 1992, "Ratchetting in Cyclic Plasticity, Part II: Multiaxial Behavior," *International Journal of Plasticity,* Vol. 8, pp. 117–146.

Inoue, T., Ohno, N., Suzuki, A., and Igari, T., 1989, "Evaluation of Inelastic Constitutive Models under Plastic-Creep Interaction for $2\frac{1}{4}$ Cr-1Mo Steel at 600°C," *Nuclear Engineering and Design,* Vol. 114, pp. 295–309.

Iwan, W. D., 1967, "On a Class of Models for the Yielding Behaviour of Continuous and Composite Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 612–617.

Jhansale, H. R., 1975, "A New Parameter for the Hysteric Stress-Strain Behavior of Metals," ASME *Journal of Engineering Materials and Technology,* Vol. 97, pp. 33–38.

Jiang, Y., 1993, "Cyclic Plasticity with an Emphasis on Ratchetting," Ph.D Dissertation, Department of Mechanical Engineering, University of Illinois at Urbana-Champaign.

Jiang, Y., and Sehitoglu, H., 1994a, "Cyclic Ratchetting of 1070 Steel under Multiaxial Stress State," *International Journal of Plasticity,* Vol. 10, No. 5, pp. 579–608.

Jiang, Y., and Sehitoglu, H., 1994b, "Multiaxial Cyclic Ratchetting under Multiple Step Loading," *International Journal of Plasticity,* Vol. 10, No. 8, pp. 849–870.

Lamba, H. S., and Sidebottom, O. M., 1978, "Cyclic Plasticity for Nonproportional Paths, Part II: Comparison with Predictions of Three Incremental Plasticity Models," ASME *Journal of Engineering Materials and Technology,* Vol. 100, pp. 104–111.

McDowell, D. L., 1983, "Transient Nonproportional Cyclic Plasticity," Ph.D Dissertation, Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign.

McDowell, D. L., 1985a, "A Two Surface Model for Transient Nonproportional Cyclic Plasticity, Part I: Development of Appropriate Equations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 298–302.

McDowell, D. L., 1985b, "A Two Surface Model for Transient Nonproportional Cyclic Plasticity, Part II: Comparison of Theory with Experiments," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 303–308.

McDowell, D. L., 1985c, "An Experimental Study on the Structure of Constitutive Equations for Nonproportional Cyclic Plasticity," ASME *Journal of Engineering Materials and Technology,* Vol. 107, pp. 307–315.

McDowell, D. L., 1992, "Description of Nonproportional Cyclic Ratchetting Behavior," *Proc. MECAMT'92, Int. Sem. on Multiaxial Plasticity,* Cachan, France, Sept. 1–4.

Mroz, Z., 1967, "On the Description of Anisotropic Workhardening," *Journal of Mechanics and Physics of Solids,* Vol. 15, No. 3, pp. 163–175.

Mroz, Z., 1981, "On Generalized Kinematic Hardening Rule with Memory of Maximum Prestress," *Journal De Mechnique Appliqué,* Vol. 5, pp. 241–260.

Mughrabi, H., 1978, "The Cyclic Hardening and Saturation Behaviour of Copper Single Crystals," *Materials Science and Engineering,* Vol. 33, pp. 207–223.

Ohno, N., 1982, "A Constitutive Model of Cyclic Plasticity with a Nonhardening Strain Region," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 721–727.

Ohno, N., and Kachi, Y., 1986, "A Constitutive Model of Cyclic Plasticity for Nonlinear Hardening Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 395–403.

Ohno, N., and Wang, J.-D., 1991, "Nonlinear Kinematic Hardening Rule: Proposition and Application to Ratchetting Problems," *Structural Mechanics in Reactor Technology,* Transactions of the 11th International Conference on Structural Mechanics in Reactor Technology, Shibata, Ed., Vol. L, Tokyo, Japan, L22/1, pp. 481–486.

Ohno, N., and Wang, J.-D., 1993a, "Kinematic Hardening Rules with Critical State of Dynamic Recovery: Part I—Formulation and Basic Features for Ratchetting Behavior," *International Journal of Plasticity,* Vol. 9, pp. 375–390.

Ohno, N., and Wang, J.-D., 1993b, "Kinematic Hardening Rules with Critical State of Dynamic Recovery, Part II: Application to Experiments of Ratchetting Behavior," *International Journal of Plasticity,* Vol. 9, pp. 391–403.

Slavik, D., and Sehitoglu, H., 1986, "Constitutive Models Suitable for Thermal Loading," ASME *Journal of Engineering Materials and Technology,* Vol. 108, pp. 303–312.

Winter, A. T., 1974, "A Model for the Fatigue of Copper at Low Plastic Strain Amplitudes," *Philosophy Magazine,* Vol. 30, pp. 719–738.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Modeling of Cyclic Ratchetting Plasticity, Part II: Comparison of Model Simulations With Experiments

Y. Jiang[1]
Research Assistant.

H. Sehitoglu
Professor.

Department of Mechanical and Industrial Engineering,
University of Illinois at Urbana-Champaign,
1206 W. Green Street,
Urbana, IL 61801

The material constants of the new plasticity model proposed in the first part of the paper can be divided into two independent groups. The first group, $c^{(i)}$ and $r^{(i)}$ (i = 1, 2, . . . , M), describes balanced loading and the second group, $\chi^{(i)}$ (i = 1, 2, . . . , M), characterizes unbalanced loading. We define balanced loading as the case when a virgin material initially isotropic will undergo no ratchetting and/or mean stress relaxation, and unbalanced loading as the loading under which a virgin material initially isotropic will produce strain ratchetting and/or mean stress relaxation. The independence of the two groups of material constants and the interpretation of the model with a limiting surface concept facilitated the determination of material constants. We describe in detail a computational procedure to determine the material constants in the models from simple uniaxial experiments. The theoretical predictions obtained by using the new plasticity model are compared with a number of multiple step ratchetting experiments under both uniaxial and biaxial tension-torsion loading. In multiple step experiments, the mean stress and stress amplitude are varied in a stepwise fashion during the test. Very close agreements are achieved between the experimental results and the model simulations including cases of nonproportional loading. Specifically, the new model predicted long-term ratchetting rate decay more accurately than the previous models.

1 Determination of Material Constants for the New Model

In this paper, a method of establishing the material constants for the proposed model will be described, followed by direct comparison of the simulations with the experiments. The material used was a 1070 (0.7 percent carbon) steel with a pearlitic structure. For the sake of conciseness, references listed in Part I will be used but not repeated in this part, and the notation used here is consistent with Part I.

1.1 Specialization of Hardening Rule to Uniaxial Loading. In this section, we present a rather unique procedure for determining the material constants for the constitutive model outlined in Part I.

When analyzing a uniaxial tension-compression test, consider that the uniaxial loading begins from the most compressive stress state where all the backstresses are saturated (on the limiting surfaces) in the compressive direction. Consideration of the one-dimensional problem results in the following simplifications:

$$\|\boldsymbol{\alpha}^{(i)}\| = \sqrt{\tfrac{3}{2}}|\alpha_{11}^{(i)}|; \quad n_{11} = \sqrt{\tfrac{2}{3}}; \quad dp = \sqrt{\tfrac{3}{2}}d(\Delta\epsilon^p). \tag{1}$$

$\Delta\epsilon^p$ in Eq. (1) is the axial plastic strain range measured from the strain state corresponding to the most compressive stress state. Simplifying the hardening rule for the one-dimensional problem, we have the following differential equation:

$$\frac{dy}{dx} = 1 - |y|^m y \tag{2}$$

where

$$x = \sqrt{\tfrac{3}{2}}c^{(i)}\Delta\epsilon^p; \quad y = \frac{\sqrt{\frac{3}{2}}\alpha_{11}^{(i)}}{r^{(i)}}; \quad \text{and} \quad m = \chi^{(i)}. \tag{3}$$

The initial condition corresponding to the most compressive stress state is $\alpha_{11}^{(i)} = -\sqrt{\frac{2}{3}}r^{(i)}$, $\Delta\epsilon^p = 0$, $y|_{x=0} = -1$. The solution to Eq. (2) corresponding to m = 0, 2, and $+\infty$ are shown in Fig. 1. We note from Fig. 1 that all solutions to Eq. (2) are asymptotic to y = 1 and they do not have a strong dependence on the exponent m or $\chi^{(i)}$. The asymptotic curve in Fig. 1 represents a limiting surface for the one-dimensional problem such that

$$\frac{\|\boldsymbol{\alpha}^{(i)}\|}{r^{(i)}} = 1 \quad \text{when} \quad \sqrt{\tfrac{3}{2}}c^{(i)}\Delta\epsilon_{(i)}^p = 2 \quad (i = 1, 2, \ldots, M), \tag{4}$$

where $\Delta\epsilon_{(i)}^p$ is the plastic strain range at which the ith backstress $\boldsymbol{\alpha}^{(i)}$ is saturated. This relationship will be vital when attempting to determine $c^{(i)}$ and $r^{(i)}$ in subsequent discussions.

While $c^{(i)}$ (i = 1, 2, . . . , M) and k, the yield stress, may be used as variables to model transient behavior, it is assumed that they remain constant for a given reversal. Assume that $\chi^{(i)}$ (i = 1, 2, . . . , M) are infinitely large, then

$$\left(\frac{\|\boldsymbol{\alpha}^{(i)}\|}{r^{(i)}}\right)^{\chi^{(i)}+1} = \begin{cases} 0 & \text{when} \quad |\boldsymbol{\alpha}^{(i)}| < r^{(i)} \\ 1 & \text{when} \quad |\boldsymbol{\alpha}^{(i)}| = r^{(i)} \end{cases} \tag{5}$$

where $\|\boldsymbol{\alpha}^{(i)}\| = r^{(i)}$ infers saturation of the ith backstress $\boldsymbol{\alpha}^{(i)}$. Since $\chi^{(i)}$ have the most pronounced influence on ratchetting

[1] Current address: Department of Mechanical Engineering, University of Reno, Reno, NV 89557.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, June 20, 1994; final revision, Apr. 28, 1995. Associate Technical Editor: W. N. Sharpe, Jr.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

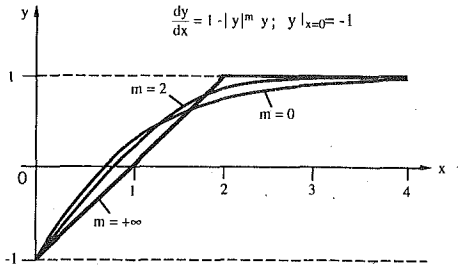


Fig. 1 Numerical solutions to the differential equation

(to be shown later), the assignment for $\chi^{(i)}$ a large number has minimal effect on the values of the constants, $c^{(i)}$ and $r^{(i)}$ (i = 1, 2, ... M). The major implication of this choice is that the uniaxial stress-strain curve simulated is piecewise linear. The plastic modulus function, when only the backstresses, j = 1, 2, ..., $(i-2)$, $(i-1)$, are saturated, can be expressed in the following form:

$$h_{(i)} = \sum_{j=i}^{M} c^{(j)} r^{(j)}. \tag{6}$$

The scalar $h_{(i)}$ represents the value of the plastic modulus function between the intervals corresponding to points $i-1$ and i.

1.2 Determination of $r^{(i)}$ from Preselected $c^{(i)}$. For the moment, if we neglect cyclic hardening and take $c^{(i)}$ (i = 1, 2, ... M) as constant, a simple procedure can be proposed to calculate the constants $c^{(i)}$ and $r^{(i)}$ (i = 1, 2, ... M) as follows. Select M points in the $\Delta\sigma - \Delta\epsilon^p$ curve from the uniaxial test (refer to Fig. 2) so that the stress range, $\Delta\sigma_{(i)}$, and plastic strain range, $\Delta\epsilon^p_{(i)}$, are known for any point i. Note that again $\Delta\epsilon^p_{(i)}$ denotes the plastic strain range at which the ith backstress $\boldsymbol{\alpha}^{(i)}$ is saturated, and $\Delta\sigma_{(i)}$ is the stress range corresponding to $\Delta\epsilon^p_{(i)}$. From Eq. (4) it follows that

$$c^{(i)} = 2\sqrt{\frac{3}{2}}\frac{1}{\Delta\epsilon^p_{(i)}} \quad (i = 1, 2, \ldots, M). \tag{7}$$

Utilizing the results of Eq. (6), the radius of the ith limiting surface can be computed as follows:

$$r^{(i)} = \frac{2}{3}\frac{H_{(i)} - H_{(i+1)}}{c^{(i)}} \quad (i = 1, 2, \ldots, M), \tag{8}$$

where

$$H_{(i)} = \frac{\Delta\sigma_{(i)} - \Delta\sigma_{(i-1)}}{\Delta\epsilon^p_{(i)} - \Delta\epsilon^p_{(i-1)}} \quad (i = 1, 2, \ldots, M), \tag{9}$$

$$\Delta\sigma_{(0)} = 2\sqrt{3}k, \quad \Delta\epsilon^p_{(0)} = 0, \quad H_{(M+1)} = 0. \tag{10}$$

The constant k is the yield stress in pure shear and the maximum stress range, $\Delta\sigma_{(M)}$, satisfies the following condition:

$$\Delta\sigma_{(M)} = 2\sqrt{\frac{3}{2}}\sum_{i=1}^{M} r^{(i)} + 2\sqrt{3}k. \tag{11}$$

$\Delta\sigma_{(M)}$ is the maximum stress range that the model is intended to simulate. The procedure (Eqs. (7)–(9)) is repeated until all $c^{(i)}$ and $r^{(i)}$ (i = 1, 2, ... M) are determined.

Within the context used to derive the current model, the cyclic hardening is considered through the coefficients $c^{(i)}$ (i = 1, 2, ... M) as functions of the accumulated plastic strain while the radii of the limiting surfaces, $r^{(i)}$ (i = 1, 2, ... M), are assumed to be constants.

1.3 Determination of $c^{(i)}$ from Preselected $r^{(i)}$. When considering cyclic hardening, the method of determining $c^{(i)}$ and $r^{(i)}$ (i = 1, 2, ... M) is somewhat more complex but we will demonstrate this as follows. If a maximum stress range is given and the yield stress has been determined, $\sum_{i=1}^{M} r^{(i)}$ can be obtained from Eq. (11) and therefore individual $r^{(i)}$ can be selected. The strategy is to find the stress and plastic strain ranges at each point in the stress-strain curve where the corresponding backstress is saturated. This procedure begins from point M and follows the order M-1, M-2, ..., 1, 0. To determine the coefficients at $(i-1)$th point (knowing ith point) the slope of the line linking the point $(i-1)$ and point i in the stress-strain reversal is calculated using the following formula:

$$H_{(i)} = \frac{3}{2}\sum_{j=i}^{M} c^{(j)} r^{(j)}. \tag{12}$$

This segment of the stress-strain curve is modeled as piecewise linear and can be described by

$$\Delta\sigma = \Delta\sigma_{(i)} + H_{(i)}(\Delta\epsilon^p - \Delta\epsilon^p_{(i)}) \quad (i = 1, 2, \ldots, M). \tag{13}$$

Point $(i-1)$ is one of the piecewise linear segment points that intersects the experimental stress-plastic strain curve, $\Delta\sigma = f(\Delta\epsilon^p)$ (Fig. 2). Once the coordinates at point $(i-1)$ are obtained, $c^{(i-1)}$ is calculated employing Eq. (7) for a known $\Delta\epsilon^p_{(i-1)}$. This procedure is repeated until all M segments have been calculated. Corresponding to a given cycle all the constants $c^{(i)}$ (i = 1, 2, ... M) can be determined. These $c^{(i)}$ (i = 1, 2, ... M) are plotted against p, the accumulated plastic strain, and then the relationship between $c^{(i)}$ (i = 1, 2, ... M) and p can be established. From the $c^{(i)} - p$ (i = 1, 2, ... M) relations, the constants $a_1^{(i)}$, $a_2^{(i)}$, $b_1^{(i)}$, and $b_2^{(i)}$ can be determined through a best fit technique. When only cyclic hardening or softening is considered, the constants $a_2^{(i)}$ and $b_2^{(i)}$ (i = 1, 2, ... M) can be set to zero and the fitting technique can be simplified.

We note that McDowell (1992) proposed a similar procedure earlier. In McDowell's method, it was implied that only one backstress part (say ith) is active while the other parts are either saturated (backstress parts less than i) or zero (backstress parts larger than i) for fully reversed uniaxial loading. It is true that when ith backstress part is active, the backstress parts $\boldsymbol{\alpha}^{(j)}$ (j = 1, 2, ... i-1) are saturated when $\chi^{(i)}$ are large numbers. However, the backstress parts $\boldsymbol{\alpha}^{(j)}$ ($j = i+1, i+2, \ldots, M$) are not zero.

1.4 Non-Masing Behavior. The non-Masing behavior is handled through the yield stress k as a function of the memory surface size, R_M. By conducting fully reversed strain-controlled uniaxial experiments at different strain amplitudes, the k-R_M relationship can be established. A clever way to establish this is by conducting an increasing/decreasing loading step test on a single specimen. In each step, the number of the loading cycles should be large enough so that the stress-strain responses stabilize. Because the non-Masing behavior is characterized by an increase in yield stress for a larger stress level, a shift of the hysteresis loops when putting the upper branches together allows determination of the change in yield stress for different stress levels. For uniaxial loading the following equation applies:

$$R_M = \sqrt{\frac{3}{2}}(\sigma_{\max} - \sqrt{3}k), \tag{14}$$

where $\sigma_{\max}$ is the maximum stress in the stabilized stress-strain hysteresis loop. From several selected stress/strain levels and their associated values of k, the k-R_M relationship can be established. The constants, k_0, a_k, and b_k, can be obtained using a best fit technique.

1.5 Memory Surface. Consider two steps from a decreasing step test. Since the stress amplitude in the second step is smaller than that of the first step, during the second step loading,

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

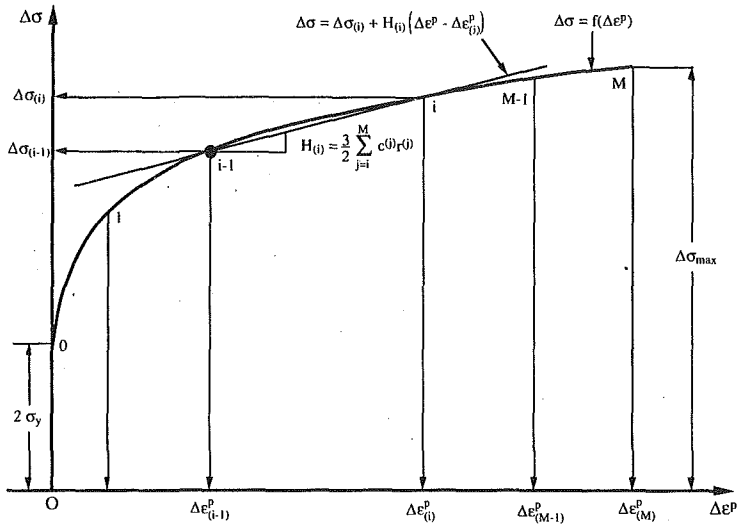


Fig. 2 Illustration of the procedure to determine $c^{(i)}$ and $r^{(i)}$ from a uniaxial stress-plastic strain reversal

$H(g) = 0$, and $|\alpha| \leq R_M$. Integrating dR_M expression results in

$$R_M = R_{M0} - c_M \int_0^{\Delta p} \left(1 - \frac{\|\alpha\|}{R_M}\right) dp, \qquad (15)$$

where Δp is the accumulated plastic strain with reference to the beginning of the second loading step. R_{M0} is the stabilized memory surface size from the first step. It is assumed that during the second loading step, there is no more cyclic hardening/softening behavior. Therefore, any change in the stress response is attributed to changes in the yield strength which are related to changes in the size of the memory surface, R_M. The instantaneous memory surface size, R_M, is calculated using Eq. (15). The constant c_M can be then determined by the trial and error fitting of R_M-Δp relation.

1.6 Determination of Exponent $\chi^{(i)}$. At this point all constants for the new model except $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) have been established. The exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) are determined from ratchetting tests. For a given stress level, the memory surface size, R_M, is a constant, and hence $\chi^{(i)}$ is a constant. The coefficients $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) can be determined from a uniaxial ratchetting test for a given stress range and mean stress. A change in the range and/or mean stress will alter the coefficients $\chi^{(i)}$ ($i = 1, 2, \ldots, M$). Using the coefficients $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) obtained for several different one-step tests at different stress levels, the constants $\chi_0^{(i)}$, a_χ, and b_χ can be determined using trial and error fitting.

1.7 Discussion of the Model Constants. The model constants for 1070 Steel are given in Table 1. The determination of the material constants $c^{(i)}$ and $r^{(i)}$ ($i = 1, 2, \ldots, M$) was based on the assumption that the exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) were large enough so that the plastic modulus function can be treated as a step function in terms of the plastic strain. In fact, the exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) have little influence on the stress-strain prediction for fully reversed tension-compression. Figure 3 proves this point with simulations using the new model for $\chi^{(i)} = 0$ and $\chi^{(i)} = +\infty$ ($i = 1, 2, \ldots, M$) along with the experimental result. When $\chi^{(i)} = 0$ ($i = 1, 2, \ldots, M$), the predicted stress-strain loop is smooth. In either case, the stress-strain simulations obtained by using the new plasticity model are in close agreement with the experimental data. Therefore, the material constants $c^{(i)}$ and $r^{(i)}$ ($i = 1, 2, \ldots, M$) can be determined following the procedures introduced previously. Further analyses indicate that the exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) have minimal influence on the predicted results for balanced loading. As will be discussed in the next section, it is unbalanced loading where the exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) play an important role. This suggests that the materials constants involved in the new plasticity model can be divided into two independent groups. One group, $c^{(i)}$ and $r^{(i)}$ ($i = 1, 2, \ldots, M$), describe balanced loading and the other group, $\chi^{(i)}$ ($i = 1, 2, \ldots, M$), characterizes unbalanced loading. The independence of the two groups of material constants facilitates the determination of material constants. In view of the common features associated with the Armstrong-Frederick type models, the procedures outlined above for determining the material con-

Table 1 Material constants used in the new plasticity model for 1070 steel ($M = 10$)

$c^{(i)}$	$c_0^{(1)} = 1510$ $c_0^{(2)} = 461$ $c_0^{(3)} = 177$ $c_0^{(4)} = 77$ $c_0^{(5)} = 39$ $c_0^{(6)} = 20$ $c_0^{(7)} = 12$ $c_0^{(8)} = 6.7$ $c_0^{(9)} = 4.8$ $c_0^{(10)} = 2.7$ $a_1^{(i)} = a_2^{(i)} = 0$ (i=1, 2, ..., 10)
$r^{(i)}$	$r^{(1)} = r^{(2)} = r^{(3)} = r^{(4)} = r^{(5)} = r^{(6)} = r^{(7)} = r^{(8)} = r^{(9)} = 63.5$ MPa $r^{(10)} = 245$MPa
$\chi^{(i)}$	$a_\chi = 0.0293$ $b_\chi = 0.0128$ MPa^{-1} $Q_0^{(1)} = 0.25$ $Q_0^{(2)} = 1.0$ $Q_0^{(3)} = Q_0^{(4)} = 2.3$ $Q_0^{(5)} = Q_0^{(6)} = Q_0^{(7)} = Q_0^{(8)} = Q_0^{(9)} = 2.75$ $Q_0^{(10)} = 4.5$
Others	$c_M = 10$ $k_1 = 92.7$ MPa $a_k = 0.0361$ $c_k = 0.0094$ MPa^{-1}

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

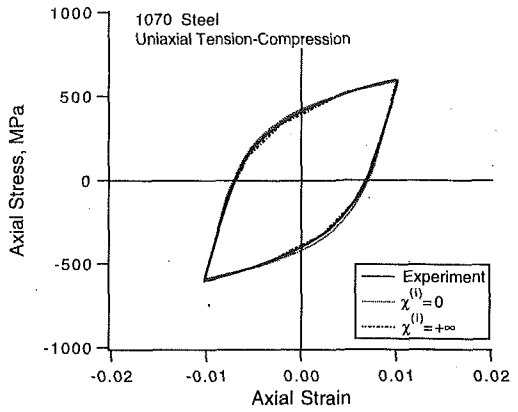


Fig. 3 Demonstration of the minimal influence of $\chi^{(i)}$ on the stress-strain description of a uniaxial balanced loading

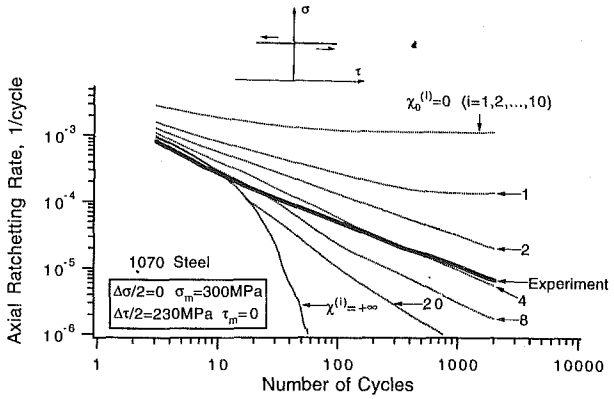


Fig. 5 Relationship between $\chi^{(i)}$ and ratchetting rate predicted by the new model for nonproportional axial-torsion consisting of alternating shear with constant axial stress

stants can be also applied to the original Armstrong-Frederick model (1966), the model of Chaboche et al. (1979), and the Ohno-Wang model (1993).

2 Experimental Procedure and Simulations

The material used in this investigation is a 1070 steel which displays a pearlitic microstructure. All data acquisition and control were performed with microcomputers. Two kinds of tests were conducted on both uniaxial solid specimens and biaxial tubular specimens. One is a single step test in which the stress magnitudes are constant during the test. The other is a multiple step test, which is composed of several single step tests. The details of material can be found in an earlier work by Jiang and Sehitoglu (1994a,b) and early references on this class of materials can be found in Slavik and Sehitoglu (1987).

Experimental data for 1070 steel (Jiang, 1993; Jiang and Sehitoglu, 1994a,b) are utilized to examine the new model's capability to predict ratchetting. The uniaxial experiments were conducted using solid circular specimens. The axial-torsion biaxial experiments were conducted employing thin-walled tubular specimens. In the simulations using the new plasticity model, the number of terms for the backstress expansion, M, was 10. Cyclic hardening for 1070 steel is small and is neglected. In Figs. 4 and 5 the experimental ratchetting rates are presented on logarithmic coordinates along with simulations with different values for $\chi^{(i)}$ (i = 1, 2, . . . , 10). The solid lines in Figs. 4 and 5 are obtained from a uniaxial experiment and a nonproportional axial-torsion experiment, respectively. The nonproportional axial-torsion experiment consist of fully reversed shear with superimposed static axial stress. In Fig. 5 the ratchetting in the shear direction is not shown because in this case both the experiment and the predictions demonstrate no shear ratchetting. In both cases, the experimental axial ratchetting rate varies from around 10^{-3} per cycle at the beginning of the loading to on the order of 10^{-6} per cycle after a few thousand cycles. Since the experimental results are approximately a straight line on log-log coordinates, the ratchetting rates follow a power law relation with respect to the loading cycles. Figures 4 and 5 reveal the pronounced influence of the exponents $\chi^{(i)}$ on the ratchetting rate predicted for both proportional and nonproportional loading. The larger the numerical values of the exponents $\chi^{(i)}$ (i = 1, 2, . . . , 10), the faster the models predict ratchetting rate decay. This contrasts with the apparent insensitivity of $\chi^{(i)}$ (i = 1, 2, . . . , 10) previously discussed for the fully reversed strain-controlled uniaxial loading conditions (Fig. 3). When $\chi^{(i)} = 0$ (i = 1, 2, . . . , 10) the model predicts a slight ratchetting rate decay for the first 100 cycles and then constant ratchetting. When $\chi^{(i)} = +\infty$ (i = 1, 2, . . . , 10), no ratchetting is predicted for the uniaxial loading. This is because the model with $\chi^{(i)} = +\infty$ (i = 1, 2, . . . , 10) produces a perfect hysteresis loop closure. For the nonproportional loading consisting of fully reversed shear with constant axial stress (Fig. 5), the model with $\chi^{(i)} = +\infty$ (i = 1, 2, . . . , 10) predicts ratchetting rate decay for about 60 cycles before ratchetting arrest. This confirms that, in the absence of cyclic hardening, the new plasticity model can predict (i) long-term ratchetting rate decay and (ii) ratchetting from near constant rate to zero rate (ratchetting arrest) for both proportional and nonproportional loading.

The exponents $\chi^{(i)}$ (i = 1, 2, . . . , M) in the new model are obtained from the analysis of the single-step uniaxial experimental ratchetting results, and the constants in the expression for $\chi^{(i)}$ (i = 1, 2, . . . , M) are determined via the procedures forwarded previously. These constants are listed in Table 1 for 1070 Steel. The ratchetting predictions using these constants are compared with the experimental results in Figs. 6 through 8. In all the figures solid

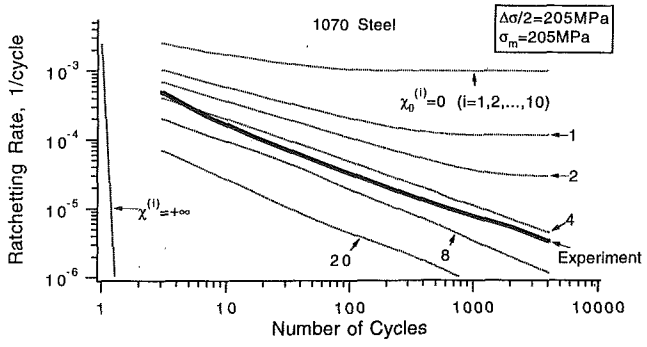


Fig. 4 Relationship between $\chi^{(i)}$ and ratchetting rate predicted by the new model for uniaxial loading

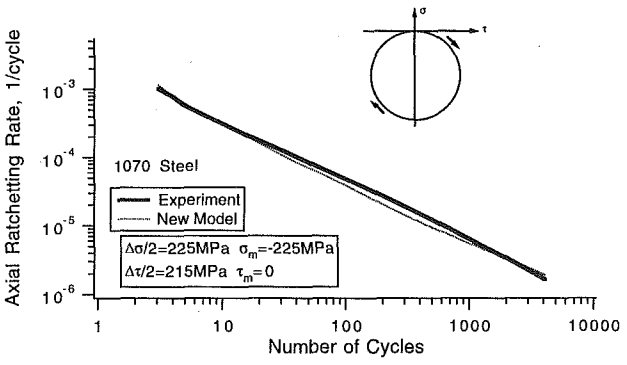


Fig. 6 Comparison of experimental data and axial ratchetting rate predicted by the new model for an "ellipse" shaped axial-torsion loading path (both experiment and new model simulation show no shear ratchetting)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

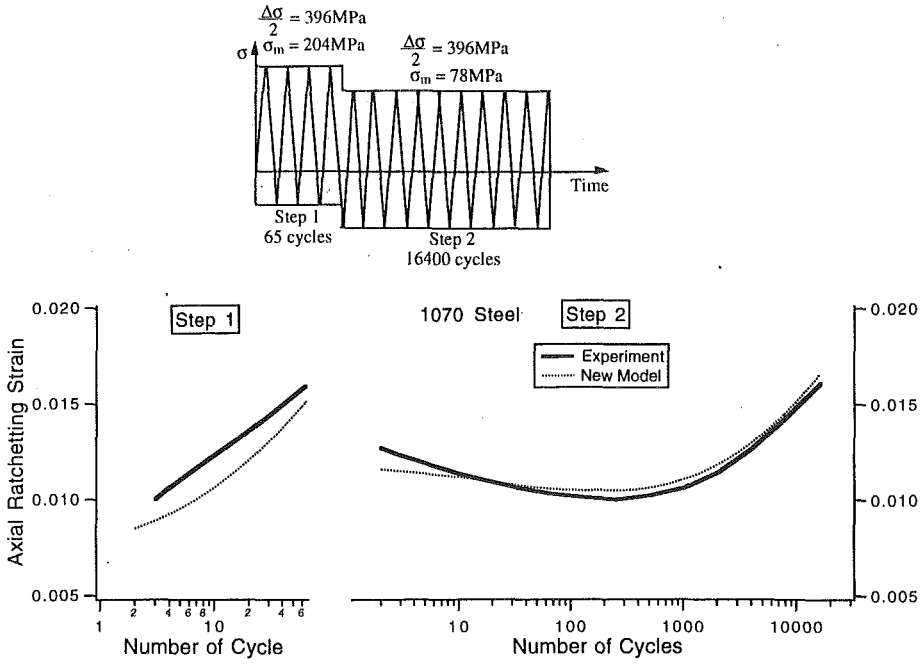


Fig. 7 Comparison of experimental data and ratchetting strain predicted by the new model for a two-step uniaxial loading

lines represent experimental results and dotted lines are predictions obtained using the new plasticity model. Shown in Fig. 6 is the comparison of experimental data and ratchetting rate predicted by the new model for an "ellipse" shaped axial-torsion loading path. The axial ratchetting rate shown is the absolute value because the actual axial ratchetting is in the compressive axial direction. Clearly, the agreement between experiment and simulation is very close for the long loading history. The ratchetting in shear direction is not

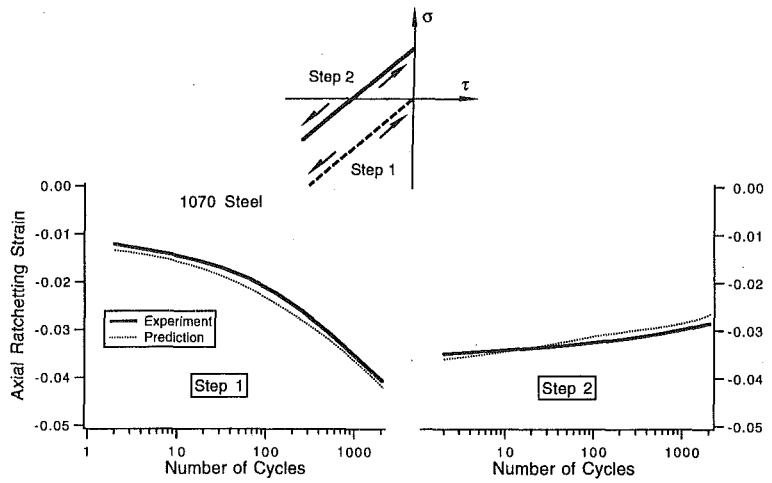


Fig. 8(*a*)

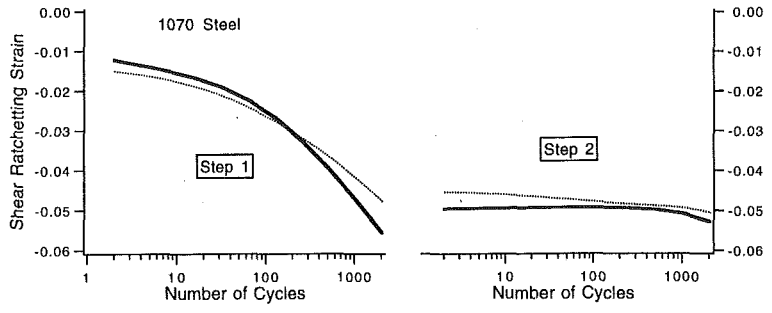


Fig. 8(*b*)

Fig. 8 Comparison of experimental data and ratchetting strain predicted by the new model for a two-step nonproportional loading history; (*a*) ratchetting in the axial direction, (*b*) ratchetting in the shear direction

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

shown since both the experiment and simulation produces no ratchetting in shear direction.

Figure 7 is a two-step uniaxial loading case. Because of a reduction of mean stress, there is a ratchetting direction change in Step 2. Obviously, the model correlates with the experiment very well. It should be noted that in this case ratchetting strain instead of ratchetting rate is used. The difference between experiment and simulation for Step 1 is mainly contributed by the first few cycles. Notably in Step 2 the model mimics the experimental ratchetting direction change appropriately and the number of loading cycles is up to 2×10^4.

Comparison of experimental data and ratchetting strain predicted by the new model for a two-step loading history is displayed in Fig. 8. Step 1 of this loading history is a axial-torsion proportional loading. In Step 1 there is a compressive mean stress in axial direction and a mean shear stress. In Step 2 the mean axial stress is reduced to zero. Step 2 is nonproportional because the ratio of axial to shear stress changes during the cycle. The axial strain component is shown in Fig. 8(a) and the shear component is presented in Fig. 8(b). Under Step 1, ratchetting rate decays in both axial and shear directions. During Step 2 loading, the axial ratchetting direction is opposite to that of Step 1 and the shear ratchetting rate decreases considerably. From Fig. 8(a), we can find that the model duplicates the experimental axial ratchetting behavior very well. In the shear direction as shown in Fig. 8(b), the agreement between experiment and simulation is not as close as that in the axial direction. This is partly because the shear stress was assumed to be homogeneous despite a small gradient in the radial direction of the thin walled cylinder.

The number of terms for the backstress expansion, M, has a weak influence on the predicted stress-strain responses for balanced loading, but a strong influence on the ratchetting prediction for unbalanced loading (Jiang, 1993). It should be noted that the selection of $M = 10$ in the previous simulations does not imply that the model inherently requires a large number of backstress terms. From the aforementioned discussion, for ratchetting rate predictions, $M = 5 \sim 10$ seems sufficient. Computer simulations for a large number of loading cycles is time consuming when both M and the number of incremental steps for a loading cycle increase. Compromises have to be made between the accuracy of the prediction and the number of loading cycles simulated. It is suggested that detailed ratchetting predictions be made for a limited number of loading cycles, and the results be extrapolated for longer loading histories using a power relation between the ratchetting rate and number of loading cycles (Jiang, 1993).

For single step loading, each $\chi^{(i)}$ has a significant influence on the predicted ratchetting rate for a limited range of loading cycles. The exponents $\chi^{(i)}$ with small i's have strong control over the initial ratchetting rate, and a weak influence on the long term ratchetting. The exponents $\chi^{(i)}$ with large i's on the other hand control the ratchetting predicted for a large number of loading cycles at a given stress level. The relationship between the ratchetting rate predicted and the selection of the exponents $\chi^{(i)}$ illustrates how the exponents $\chi^{(i)}$ ($i = 1, 2, \ldots, M$) would be selected. Detailed discussions can be found in Jiang (1993).

3 Further Discussion on Ratchetting

We attempt to gain further insight into the ratchetting phenomenon here. Deriving the ratchetting increment per cycle in closed form for nonproportional loading is virtually impossible. However, it may be helpful to discuss a cyclically stable unbalanced uniaxial simulation. An arbitrary stress-strain response is schematically shown in Fig. 9 for a uniaxial loading. Point 1 denotes the beginning of the cycle, and the material is stressed in the forward direction to 2. At point 2, the material is unloaded until the stress 3 is reached.

The ratchetting rate (ratchetting strain per cycle) for this case is

$$\frac{d\epsilon_r}{dN} = \sum_{i=1}^{M} \frac{d\epsilon_r^{(i)}}{dN} = \int_1^3 d\epsilon^p \tag{16}$$

where

$$\frac{d\epsilon_r^{(i)}}{dN} = \frac{c^{(i)} r^{(i)} \left[\int_1^2 \left(\frac{\|\boldsymbol{\alpha}^{(i)}\|}{r^{(i)}} \right)^{\chi^{(i)}+1} \mathbf{L}^{(i)} : \mathbf{n} d\epsilon^p + \int_2^3 \left(\frac{|\boldsymbol{\alpha}^{(i)}|}{r^{(i)}} \right)^{\chi^{(i)}+1} \mathbf{L}^{(i)} : \mathbf{n} d\epsilon^p \right]}{\sum_{j=1}^{M} c^{(j)} r^{(j)}} \quad (i = 1, 2, \ldots, M). \tag{17}$$

The term, $d\epsilon_r^{(i)}/dN$, in the previous equations represents the contribution to ratchetting rate by the variation of an individual backstress. A few insights can be gained from the examination of Eq. (17). The denominator on the right side of Eq. (17) is a constant since a cyclically stable material was assumed. If $\boldsymbol{\alpha}^{(i)}$ varies symmetrically (i.e., no mean value), the integral on the right-hand side of Eq. (17) will be zero, which results in $d\epsilon_r^{(i)}/dN = 0$. The quantity $(\|\boldsymbol{\alpha}^{(i)}\| r^{(i)})^{\chi^{(i)}+1}$ is equal to or less than 1.0. Therefore, if $\chi^{(i)} = +\infty$, then $d\epsilon_r^{(i)}/dN = 0$. For balanced loading, the scalar product of vector, $\mathbf{L}^{(i)}$ and $\mathbf{n}$ will change sign for loading from 1 to 2 and 2 to 3 (refer to Fig. 9) and results in the sum of these integrals to be approximately zero, i.e., $d\epsilon_r^{(i)}/dN \approx 0$. For a single step, fully reversed uniaxial loading, all the backstresses vary symmetrically and no ratchetting is predicted.

Figure 10 shows a two-step unbalanced uniaxial loading. For the one-dimensional problem of tension-compression, $\boldsymbol{\alpha}^{(i)}$ can be represented by the value of the component $\alpha_{11}^{(i)}$. The term, The material constants employed are not listed because $\sqrt{\frac{3}{2}}\,\alpha_{11}^{(i)}/r^{(i)}$, is a normalized quality representing the backstress $\boldsymbol{\alpha}^{(i)}$ for the uniaxial case. In Fig. 10 the number of terms for the backstress, M, is 5 since the desire is to qualitatively display the behavior of $\boldsymbol{\alpha}^{(i)}$. The selection of the material constants does not qualitatively alter the ensuing points of discussion. The predicted ratchetting rates for both steps are shown in Fig. 10(b), and the variations of the backstresses with the loading history are presented in Fig. 10(c). During Step 1 loading, ratchetting rate decay is predicted. The corresponding variations

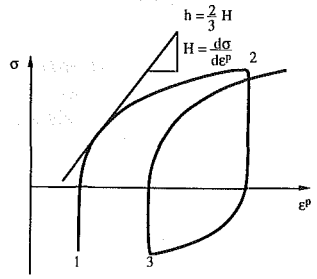


Fig. 9 Schematic representation of strain ratchetting for a uniaxial simulation

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

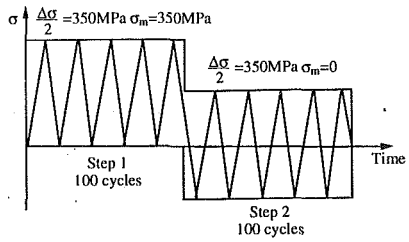


Fig. 10(*a*)

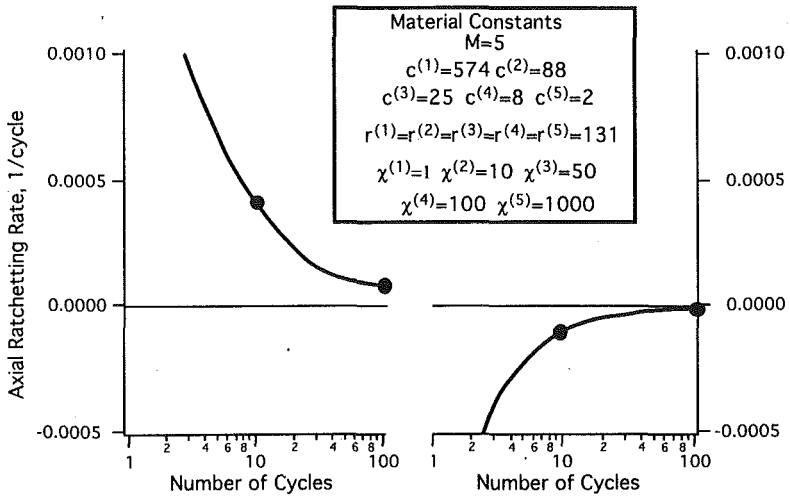


Fig. 10(*b*)

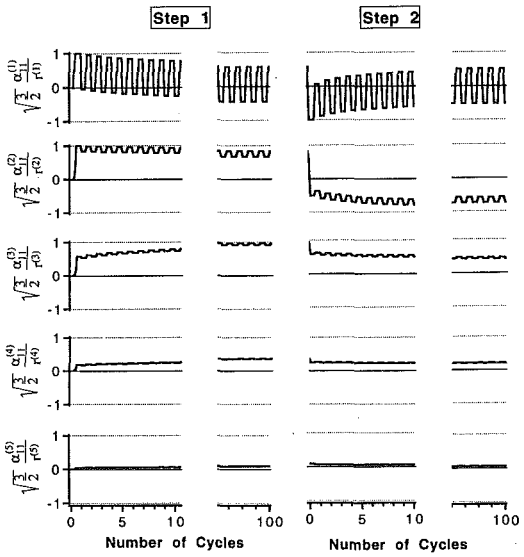


Fig. 10(*c*)

Fig. 10 Demonstration of ratchetting rate decay predicted by the new model for a two-step uniaxial loading; (*a*) two-step uniaxial loading, (*b*) ratchetting rate, (*c*) $\alpha_{11}^{(i)}$ variations with loading cycles

of the backstresses shown on the left side of Fig. 10(*c*) are responsible for the ratchetting rate decay. The magnitudes of backstress variations are in an order consistent with the size of $c^{(i)}$ (i = 1, 2, 3, 4, 5). Subsequently, the mean values of $\alpha_{11}^{(1)}$ and $\alpha_{11}^{(2)}$ saturate to zero with increasing cycles and the mean values of $\alpha_{11}^{(4)}$ and $\alpha_{11}^{(5)}$ increase at the same time. The amplitudes of the backstresses do not change with loading cycles. According to Eq. (17), for the same variation of the backstress and $\chi^{(i)}$, a larger value of $c^{(i)}r^{(i)}$ will result in a larger ratchetting rate. Because it was assumed that $r^{(1)} = r^{(2)} = r^{(3)} = r^{(4)} = r^{(5)}$, a larger $c^{(i)}$ results in a larger $c^{(i)}r^{(i)}$. The decrease of the mean values of $\alpha_{11}^{(1)}$ and $\alpha_{11}^{(2)}$ (corresponding to larger $c^{(i)}r^{(i)}$ values) mainly contributes to the initial ratchetting rate decay predicted. It should be noted that the mean value of the total backstress is proportional to the mean stress of the stress-controlled uniaxial loading cycle. Therefore, the sum of all the backstress has a fixed mean value for a given step in the loading. The decrease in the mean values of $\alpha_{11}^{(1)}$ and $\alpha_{11}^{(2)}$ is consistent with the increase of the mean values of $\alpha_{11}^{(4)}$ and $\alpha_{11}^{(5)}$. It is this shifting of the mean values of the backstresses in combination with the smaller amplitude of $\alpha_{11}^{(4)}$ and $\alpha_{11}^{(5)}$ that contributes to the continued ratchetting rate decay.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

When the mean stress is zero in Step 2, the model predicts ratchetting in the direction opposite to that in the first step. From the right side of Fig. 10(*b*), the ratchetting rate (absolute value) also decreases with increasing number of loading cycles, which again can be explained by the variations of the backstresses (refer to the right side of Fig. 10(*c*)). When Step 2 with a different mean stress begins, there are sudden changes in the backstresses, which corresponds to a change in the ratchetting rate. The backstresses $\alpha_{11}^{(1)}$ and $\alpha_{11}^{(2)}$ with initial negative mean levels contribute to the ratchetting rate in the negative direction, while the other three backstress parts with positive mean values produce ratchetting in the positive direction. As was previously stipulated, a larger amplitude of a backstress contributes more to the ratchetting rate, hence the overall ratchetting rate is negative. Since the mean stress in Step 2 is zero, all the mean values of the backstresses will approach to zero with increasing number of loading cycles. As a result, the ratchetting rate decreases with increasing number of cycles. For ratchetting under nonproportional loading, the ratchetting predicted using the new model may be explained in a similar way. However, for nonproportional loading, the ratchetting rate is not only dependent on the variations of the backstresses but also on the variations of $\mathbf{L}^{(i)} : \mathbf{n}$ ($i = 1, 2, \ldots, M$). A straightforward illustration is more difficult for unbalanced nonproportional loading.

4 Conclusions

The following conclusions are drawn from the work:

1 The new plasticity model was applied to the ratchetting predictions of 1070 steel. Its capability to improve long-term ratchetting and multiple step ratchetting predictions was demonstrated.

2 The material constants in the new plasticity model can be divided into two independent groups; one group, $c^{(i)}$ and $r^{(i)}$ ($i = 1, 2, \ldots, M$), which describes balanced loading and the other group, $\chi^{(i)}$ ($i = 1, 2, \ldots, M$), which characterizes unbalanced loading. The independence of the two groups of material constants and the interpretation of the model with a limiting surface concept facilitated the determination of material constants.

3 The material displays minimal cyclic hardening, however, non-Masing behavior is observed. The model correlated the non-Masing behavior accurately.

4 For multiple step loading, the material exhibits a memory of the previous loading history. Due to this memory effect the strain ratchetting may accelerate under certain circumstances, although such an acceleration in ratchetting rate is short-lived. The model predicts the prior loading history effect on the subsequent ratchetting, the change in ratchetting direction and the dissipation of this effect with increasing number of loading cycles.

Acknowledgments

Professor Sehitoglu acknowledges the financial support from Department of Energy, Basic Energy Sciences, DOE DEFG02-93ER14393. The continued support of Association of American Railroads, Technical Center, Chicago, IL and the encouragement of Dr. Dan Stone, AAR, are also acknowledged.

References

Slavik, D., and Sehitoglu, H., 1987, "A Constitutive Model for High Temperature Loading, Part I—Experimentally Based Form of the Equations," *Thermal Stress, Material Deformation, and Thermo-mechanical Fatigue,* ASME, New York, PVP-123, pp. 65–73.

Slavik, D., and Sehitoglu, H., 1987, "A Constitutive Model for High Temperature Loading, Part II—Comparisons of Simulations with Experiments," *Thermal Stress, Material Deformation, and Thermo-mechanical Fatigue,* ASME, New York, PVP-123, pp. 75–82.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part I: Elastic Response

D. L. Littlefield
Southwest Research Institute,
San Antonio, TX 78228
Mem. ASME

The effects of electromagnetic fields on instabllities in metal cylinders are considered in this analysis. The cylinder is assumed to be infinitely long and perfectly conducting. An axial electric current is introduced in the cylinder, giving rise to an azimuthal magnetic field in the surrounding vacuum, causing mechanical distortion in the cylinder. The current is assumed to be small so that the deformation remains elastic; in an accompanying paper (Littlefield, 1996) larger currents are considered where plastic flow becomes important. After solutions to the idealized motion of the cylinder under uniaxial strain conditions are developed, small perturbations to the motion are considered. The equations governing the motion of these disturbances are derived using linear perturbation theory. Solutions to the equations indicate that electromagnetic fields can have a substantial effect on the stability spectrum in the cylinder. In general, the frequency of oscillating perturbations is suppressed by the azimuthal magnetic field, and distending instabilities are possible if the magnetic field is above a threshold value. The underlying physical mechanisms contributing to these deviations are proposed.

1 Introduction

The formation and propagation of instabilities in cylindrical metal bars when exposed to large electromagnetic fields has become an area of intense research over the past few years. Arenas for utilization of this research include diverse applications such as exploding wire technology, microprocessor manufacturing, and penetrator disruption. When large electric currents are passed through thin wires, for instance, the wire has a tendency to *explode,* breaking up into a large number of thin striated segments (Chace and Moore, 1968). Although rapid expansion of the plasma formed from the vaporized wire certainly contributes to the observed breakup, electromagnetically-induced instabilities may also be partially responsible for the phenomenon. On the other hand, electromagnetic instabilities may also contribute to the failure of microprocessors and other microelectronic devices, where the large current densities along conducting paths might contribute to fatigue and creep failures. Moreover, if electromagnetic fields can be used to disrupt and break up penetrators, the magnetic field itself could be utilized as an effective medium for mitigation of threats to armored vehicles. In all of these applications, the coupling of electromagnetic effects to the mechanical behavior of the material is of paramount importance. This two-part paper examines how the fundamental instability modes in metal cylinders are altered by the presence of electromagnetic fields, and also how different instability modes are introduced.

Classical treatments considering the propagation of displacements in cylindrical bars were initiated many years ago, beginning with the work by Rayleigh (1877), who obtained long-wavelength approximations for the phase velocities of waves traveling through an infinitely long, elastic cylindrical bar. Complete solutions to this problem were first performed by Onoe et al. (1962) for the propagation of axisymmetric waves, including the bounding surfaces for various limiting conditions. A comprehensive review of solutions to stress wave propagation problems in elastic media can be found many texts, including the monograph by Graff (1975).

It has long been observed that magnetic fields can be used either to inhibit or enhance the stability of columns comprised of a conducting material. This phenomenon has been identified both in magnetic fluids and conducting solids. The classical work by Chandrasekhar (1961) demonstrated that axial magnetic fields can stabilize cylindrical fluid columns subject to capillary instabilities. The stability of capillary magnetic fluid jets when exposed to axial or azimuthal magnetic fields was examined by Malik and Singh (1991). Their analyses demonstrated that hydrodynamic cylindrical fluid columns were conditionally stable, depending on the ratio of electromagnetic to surface tension forces. Chattopadhyay and Moon (1975) studied the magnetoelastic buckling in conducting solid rods under magneto-quasi-static conditions. They found that above a critical level of current, a *kink* instability can be initiated in the rod that causes buckling. Nowacki (1990) investigated a similar problem of the vibration and stability of a rod when subjected to a self-induced azimuthal field as well as a lateral field. Littlefield (1990, 1991a, 1991b, 1994) has studied the effects of electromagnetic fields of various orientations on the stability of perfectly plastic jets. The analyses predicted that azimuthal magnetic fields always tend to enhance the growth rate of the naturally occurring instabilities, but axial magnetic fields could be stabilizing under certain conditions. Comprehensive reviews of problem-solving techniques in the general area of magnetosolid mechanics, as well as reviews of previous work, are given in the texts by Moon (1984) and Maugin (1988).

When large electric currents are suddenly introduced into conducting rods, the current remains on the surface for a short time, then eventually diffuses to the interior of the rod. The magnetic Reynolds number is the dimensionless parameter that measures the relative importance of current diffusion to the specific problem; when the magnetic Reynolds is very large, the current can be assumed to remain on the surface of the rod. The abrupt introduction of current to a conducting cylinder is the problem of interest in this investigation, where the magnetic

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Dec. 19, 1994; final revision, Oct. 20, 1995. Associate Technical Editor: W. J. Drugan.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Reynolds number can be quite large. This is distinguished from the aforementioned studies of instabilities in conducting elastic rods, where current diffusion occurs on a time scale much faster than the mechanical time scale, and a magneto-quasi-static analysis is appropriate.

In this analysis the stability of an infinitely long, perfectly conducting metal cylinder is considered. An axial electric current is introduced, creating an azimuthal magnetic field in the surrounding vacuum, and interacting with the current to generate an electromagnetic *pinching* force, directed towards the axis of the cylinder. This force gives rise to a condition of uniform uniaxial strain in the cylinder. Two-dimensional axisymmetric perturbations to this motion are considered, and the equations governing the motion of these disturbances are derived using linear perturbation theory. These equations are solved to determine how the amplitude of small disturbances changes over time.

Several assumptions have been made to make the analysis tractable, yet to retain as realistic conditions as possible. The cylinder is assumed to be infinitely long and perfectly conducting. This implies that end effects are neglected and that no diffusion of the magnetic field occurs to the interior of the cylinder. Although the assumption on conductivity may not be realistic under many conditions, other analyses for perfectly plastic jets (Littlefield and Powell, 1990; Littlefield, 1991a; Littlefield, 1994) have demonstrated that the assumption of finite conductivity does not alter substantially the qualitative mechanical behavior of instabilities; the principal effect is to introduce Joule heating and spatially dependent mechanical softening in the jet. Thus, it seems reasonable to expect that the stability of a cylinder with finite electrical conductivity will be qualitatively similar to the perfectly conducting case. Deformation in the cylinder is limited to the elastic response, and the elastic moduli are further assumed to be constant. In an accompanying paper (Littlefield, 1996), this assumption is relaxed and elastic-plastic flow is considered. Both the idealized and perturbed motion of the cylinder are limited to small strains; however, geometric nonlinearities associated with the electromechanical coupling are retained by evaluating Lorentz forces in the deformed configuration. Only axisymmetric motion of the cylinder is permitted, so azimuthal deformation of the cylinder is not considered. The cylinder is also assumed isothermal, and resistive Joule heating is neglected. Thus, energy conservation is not considered in this analysis.

The remainder of this paper is arranged as follows. In Section 2, the mathematical model is described and the basic equations applicable in the cylinder and the surrounding vacuum are formulated. These equations are solved under idealized conditions in Section 3. Small perturbations to this motion are considered in Section 4, and the equations governing the motion of these small disturbances are developed and solved. Solutions to the perturbation equations are shown in Section 5, and the underlying physical mechanisms responsible for the instabilities observed are elucidated. Finally, in Section 6 some general conclusions and comments relevant to the application of the results are presented.

2 Mathematical Model

Consider the deformation of a uniform, perfectly conducting, infinitely long cylinder as is depicted in Fig. 1. The radius at an arbitrary point along the surface of the cylinder is denoted as $r_b(z, t)$, and the outward unit normal vector at this point as $\mathbf{n}$. An axial electric current is introduced at time $t = 0$ and is assumed to be concentrated on the surface of the cylinder. This electric current gives rise to an azimuthal magnetic field in the surrounding vacuum. The magnetic field interacts with the surface current to create a Lorentz force, causing mechanical distortion to occur in the cylinder.

Let the density be given by ρ, the radial and axial displacement by u_r and u_z, respectively; the pressure by p, the deviatoric

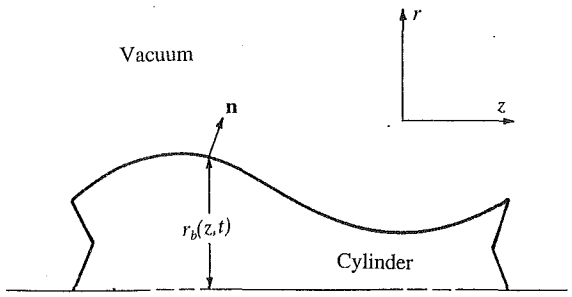


Fig. 1 Model for stability calculations

components of the stress tensor by τ_{rr}, τ_{rz}, and τ_{zz}; and the azimuthal magnetic field by B_θ. Then the governing differential equations of motion in the metal cylinder, simplified in accordance with the assumptions outlined in the introduction, are given by

$$\rho \frac{\partial^2 u_r}{\partial t^2} = -\frac{\partial p}{\partial r} + \frac{\partial \tau_{rr}}{\partial r} + \frac{2\tau_{rr} + \tau_{zz}}{r} + \frac{\partial \tau_{rz}}{\partial z}, \tag{1}$$

and

$$\rho \frac{\partial^2 u_z}{\partial t^2} = -\frac{\partial p}{\partial z} + \frac{\partial \tau_{rz}}{\partial r} + \frac{\tau_{rz}}{r} + \frac{\partial \tau_{zz}}{\partial z}, \tag{2}$$

where the pressure, deviatoric stress, and displacements are related as

$$\left.\begin{aligned} \frac{\partial u_r}{\partial r} &= \frac{\tau_{rr}}{2G} - \frac{p}{3K} \\ \frac{\partial u_r}{\partial z} + \frac{\partial u_r}{\partial r} &= \frac{\tau_{rz}}{G} \\ \frac{\partial u_z}{\partial z} &= \frac{\tau_{zz}}{2G} - \frac{p}{3K} \\ \frac{\partial u_r}{\partial r} + \frac{u_r}{r} + \frac{\partial u_z}{\partial z} &= -\frac{p}{K} \end{aligned}\right\}, \tag{3}$$

and G and K are the shear and bulk moduli, respectively. Equations (1)–(3) are the simplified forms of the momentum equation and constitutive relation for a linear elastic material, assuming the elastic strains are small. Note that the Lorentz forces do not appear in Eqs. (1)–(2) due to the assumption on conductivity. When the conductivity is very large, no diffusion of the magnetic field occurs, so if the current is assumed to be initially concentrated on the surface, it remains there, and the current, magnetic and electric fields all vanish in the interior of the cylinder. Thus, in this analysis, all mechanical effects attributed to the Lorentz force remain concentrated on the surface.

In the vacuum surrounding the cylinder, the conductivity is zero, thus Maxwell's equations can be conveniently reduced to a single equation for the azimuthal magnetic field, given as

$$\frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial}{\partial r}(rB_\theta)\right] + \frac{\partial^2 B_\theta}{\partial z^2} = 0. \tag{4}$$

The use of Eqs. (1)–(4) to determine the motion in the cylinder and the surrounding vacuum implies that certain boundary conditions must be satisfied at the interface. In particular, momentum must still be conserved at the boundary and Maxwell's equations must be satisfied. A general procedure for determining the correct boundary conditions is to integrate the general forms of the governing equations across a layer of thickness Δ positioned at the interface, then taking the limit as $\Delta \to 0$ (e.g., see Kruskal and Schwartzchild, 1954). This technique

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

insures that the governing equations will still be satisfied at the boundary. In addition, the presence of a surface current $\mathbf{J}^*$ must permitted such that the integral of the current density $\mathbf{J}$ across the boundary yields $\mathbf{J}^*$ as $\Delta \to 0$. Application of this procedure yields

$$\frac{\partial r_b}{\partial t} = v_r - \frac{\partial r_b}{\partial z} v_z, \tag{5}$$

$$-p + \tau_{rr} + \frac{B_\theta^2}{2\mu} - \frac{\partial r_b}{\partial z} \tau_{rz} = 0, \tag{6}$$

$$\frac{\partial r_b}{\partial z}\left(-p + \tau_{zz} + \frac{B_\theta^2}{2\mu}\right) - \tau_{rz} = 0, \tag{7}$$

and

$$B_\theta = \frac{\mu I}{2\pi r}, \tag{8}$$

where v_r and v_z are the radial and axial velocities, I is the constant electric current, μ is the magnetic permeability, and these conditions are all to be evaluated at $r = r_b$. The velocities are related to the displacements as

$$\frac{\partial u_r}{\partial t} = v_r, \quad \frac{\partial u_z}{\partial t} = v_z. \tag{9}$$

Together, Eqs. (1)–(9) represent a complete system of ten coupled partial differential equations with the appropriate boundary conditions that can be solved for the ten unknowns in the cylinder and vacuum: u_r, u_z, v_r, v_z, p, τ_{rr}, τ_{rz}, τ_{zz}, r_b, and B_θ. Solutions to these equations are used to determine the motion of the cylinder.

3 Idealized Motion

In the idealized steady motion of the cylinder, the uniform axial electric current gives rise to an azimuthal magnetic field, which interacts with the current to create a Lorentz force concentrated at the boundary and directed radially towards the center of the cylinder. This force creates a displacement in the radial boundary of the cylinder such that the outwardly directed radial restoring force in the cylinder is precisely balanced by the Lorentz force. Thus, the base solutions to the equations of motion are characterized by one-dimensional uniaxial strain conditions along the radius of the cylinder, such that the displacements are time invariant and the velocities are zero. If the electric current is sufficiently low in magnitude, then the deformation remains elastic and the displacements can be determined from solutions to Eqs. (1) and (3), together with the boundary conditions in Eqs. (5)–(6) and (8). Here it is appropriate to use a time-integrated form of Eq. (5) since it is the displacements that are determined. Although elastic strains in the cylinder were assumed to be small in the development of Eqs. (1) and (3), the geometric nonlinearity associated with the Lorentz force is incorporated in this analysis by evaluating the magnetic pressure at the *deformed* boundary, instead of the original boundary. Note that the inclusion of this nonlinearity is more critical than accounting for large elastic strains. Under uniaxial compression conditions, neglecting the latter generally requires $(1 + e)$ to be close to 1, whereas neglecting the former requires $(1 + e)^{-2} \approx 1$, where e is an average measure of the normalized radial elongation. With these provisions, the governing equations admit a steady-state solution for the radial displacement that is a linear function r. Thus, Eq. (3) yields for the deviatoric stress and pressure:

$$\tilde{\tau}_{rr0} = 2\Phi\tilde{u}_{rb0}/3; \quad \tilde{\tau}_{zz0} = -4\Phi\tilde{u}_{rb0}/3;$$
$$\tilde{\tau}_{rz0} = 0; \quad \tilde{p}_0 = -2\tilde{u}_{rb0}; \tag{10}$$

where $\tilde{\tau} = \tau/K$, $\tilde{p} = p/K$, and the subscript 0 is used to denote idealized quantities. $\tilde{u}_{rb0}$ is the value of the normalized displacement $\tilde{u}_{r0}$ evaluated at the boundary $r = r_{b0}$, where $\tilde{u}_{r0} = u_{r0}/r_{b0}$. With use of Eq. (10), Eq. (6) requires that $\tilde{u}_{rb0}$ satisfy

$$\tilde{u}_{rb0}^3 + 2\tilde{u}_{rb0}^2 + \tilde{u}_{rb0} + \frac{\Lambda_0}{4(1 + \Phi/3)} = 0. \tag{11}$$

The dimensionless parameters Λ_0 and Φ appearing in Eqs. (10) and (11) are defined as

$$\Lambda_0 = \frac{\mu I^2}{4\pi^2 r_{b0}^2 K}, \quad \Phi = \frac{G}{K}, \tag{12}$$

where Λ_0 physically represents a ratio of electromagnetic to compressive forces, and Φ a ratio of shear to compressive forces. Solving Eq. (11) for $\tilde{u}_{rb0}$ yields three solutions for the displacement, but only one of these is physically meaningful, and is given as

$$\tilde{u}_{r0} = \tilde{u}_{rb0}\frac{r}{r_{b0}} = -\frac{2}{3}\left\{1 - \cos\left[\frac{1}{3}\tan^{-1}\left(\frac{9}{\sqrt{3}}\frac{\sqrt{\alpha(4 - 27\alpha)}}{2 - 27\alpha}\right)\right]\right\}\frac{r}{r_{b0}}, \tag{13}$$

where $\alpha = \Lambda_0/4(1 + \Phi/3)$. Note that Eq. (13) admits real solutions for $\tilde{u}_{r0}$ only when $\alpha < \alpha_c = \frac{4}{27}$. This occurs because when $\alpha > \alpha_c$, there are no steady-state solutions for the displacement. A close examination of the boundary condition used to form Eq. (11) reveals the reason for this significant result. As the radius of cylinder is reduced, the magnetic pressure increases as the radius squared; however, the elastic restoring stress increases only linearly with radius. Thus, when the radial displacement just exceeds a certain critical value, corresponding to $\alpha = \alpha_c$, the elastic restoring stress can no longer balance the magnetic pressure, and the radial boundary continues to accelerate towards an even smaller radius. Critical values for $\tilde{u}_{rb0}$ and Λ_0 can be derived from Eq. (13) and are given as

$$\tilde{u}_{rb0c} = -\frac{1}{3}; \quad \Lambda_{0c} = \frac{16}{27}\left(1 + \frac{\Phi}{3}\right), \tag{14}$$

where the subscript c denotes the critical value. It is, of course, important to recognize that this result is derived based on the assumption of small strains in the cylinder, which is certainly violated when the cylinder is compressed down to a small radius. However, preliminary calculations (Littlefield, 1996) were performed taking into account finite elastic strains in the cylinder, and indicate that a criticality condition still exists; although the numerical values given in Eq. (14) are somewhat changed. Most importantly, however, is the fact that experiments have shown (Chace and Moore, 1968) how particularly rapid and violent the breakup of metallic wires can be when exposed to large axial currents, which seems consistent with the transcendence of some criticality condition like that imposed in Eq. (14).

If the current is large enough such that $\Lambda_0 > \Lambda_{0c}$, then transient base solutions must be considered to Eqs. (1) and (3). A calculation was performed to illustrate the behavior of the displacement beyond the critical condition. Shown in Fig. 2 is a plot of the normalized displacement at the boundary $\tilde{u}_{rb0}$ as a function of normalized time $\tilde{t}$, where $\tilde{t} = t\sqrt{K/\rho}/r_{b0}$. The values for Λ_0 and Φ were 0.8 and 0.5, respectively. It is evident from the results in the figure that the displacement at the boundary decreases quite rapidly for $\tilde{u}_{rb0} > \tilde{u}_{rb0c}$.

In the vacuum surrounding the cylinder, the solution for $B_{\theta 0}$ consistent with the uniaxial strain conditions in the cylinder, but still finite as $\tilde{r} \to \infty$, is given by

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

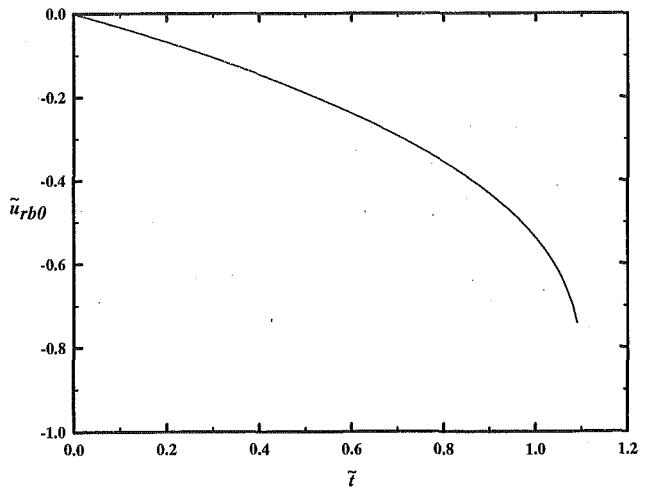


Fig. 2 Normalized displacement at the boundary $\tilde{u}_{rb0}$ versus normalized time $\tilde{t}$ for an elastic cylinder with $\Phi = \frac{1}{2}$ and $\Lambda = 1$

$$B_{\theta 0} = \frac{\mu I}{2\pi r}, \tag{15}$$

which also identically satisfies the required boundary condition given in Eq. (8).

4 First-Order Equations and Solutions

The solutions developed in the previous section are now subjected to a small axisymmetric disturbance. The equations governing the time evolution of these small perturbations to the idealized motion are derived using linear perturbation theory. These equations may be formed in the traditional manner by expanding the dependent variables in a power series of ϵ, keeping only the terms of order ϵ^1. Integration of these equations yields solutions for the time evolution of instabilities to the idealized motion of the cylinder.

If the electric current is small, then the idealized motion in the cylinder, as well as the flow of small perturbations to this motion, remains elastic. The equations governing the spatial and time dependence of these disturbances are derived using linear perturbation theory. The resulting equations are linear partial differential equations having coefficients independent of both t and z, so the axial and temporal dependences are both exponential in nature. Let

$$\left.\begin{aligned} u_{r1} &= \tilde{u}_r r_{b0} \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \\ u_{z1} &= \tilde{u}_z r_{b0} \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \\ p_1 &= \tilde{p}K \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \\ \tau_{rr1} &= \tilde{\tau}_{rr}K \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \\ \tau_{zz1} &= \tilde{\tau}_{zz}K \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \\ \tau_{rz1} &= \tilde{\tau}_{rz}K \exp(ik\tilde{z}) \exp(\omega\tilde{t}) \end{aligned}\right\}, \tag{16}$$

where the normalized axial coordinate is $\tilde{z} = z/r_{b0}$, and k and ω are the axial wave number and normalized growth rate. Then the governing equations for the first-order variables may be combined in a straightforward manner to yield a single equation for $\tilde{u}_r$, given by

$$\left\{\omega^2 + \Theta k^2 - \Theta\left(\frac{\partial^2}{\partial\tilde{r}^2} + \frac{1}{\tilde{r}}\frac{\partial}{\partial\tilde{r}} - \frac{1}{\tilde{r}^2}\right)\right\} \times \left\{\omega^2 + \Phi k^2 - \Phi\left(\frac{\partial^2}{\partial\tilde{r}^2} + \frac{1}{\tilde{r}}\frac{\partial}{\partial\tilde{r}} - \frac{1}{\tilde{r}^2}\right)\right\}\tilde{u}_r = 0, \tag{17}$$

where $\tilde{r} = r/r_{b0}$ and $\Theta = 1 + 4\Phi/3$. Physically, Θ and Φ represent normalized elastic longitudinal and transverse wavespeeds, respectively. Solutions to these equations yield the first-order radial and axial displacements in the cylinder given by

$$\tilde{u}_r = c_1 I_1(\beta_1\tilde{r}) + c_2 I_1(\beta_2\tilde{r}),$$

$$\tilde{u}_z = c_1 \frac{ik}{\beta_1} I_0(\beta_1\tilde{r}) + c_2 \frac{i\beta_2}{k} I_0(\beta_2\tilde{r}), \tag{18}$$

where $\beta_1^2 = k^2 + \omega^2/\Theta$, $\beta_2^2 = k^2 + \omega^2/\Phi$, and c_1 and c_2 are arbitrary constants.

In the vacuum surrounding the cylinder, solutions to the first-order equations for B_θ, together with the associated boundary conditions, require the perturbed magnetic field to vanish there. This result also follows directly from Ampère's Law, since the current enclosed in a cross-sectional segment of the cylinder at any axial location necessarily remains fixed.

The constants c_1 and c_2 in Eq. (18) are determined through application of boundary conditions. These boundary conditions are derived from the first-order expansions of Eqs. (5)–(8). After Eqs. (5) and (8) are used to eliminate r_b and B_θ in Eqs. (6)–(7), the two required boundary conditions may be expressed as

$$\Theta\frac{\partial\tilde{u}_r}{\partial\tilde{r}} + (\Theta - 2\Phi)\left(\frac{\tilde{u}_r}{\tilde{r}} + ik\tilde{u}_z\right) - Z\tilde{u}_r = 0, \tag{19}$$

and

$$\left\{1 - \frac{Z}{2(\Theta - \Phi)}\right\}ik\tilde{u}_r + \frac{\partial\tilde{u}_z}{\partial\tilde{r}} = 0, \tag{20}$$

where

$$Z = \frac{4(\Theta - \Phi)\Lambda}{4(\Theta - \Phi) + \Lambda}, \tag{21}$$

and these conditions are evaluated at $\tilde{r} = 1$. The parameter Λ is related to Λ_0 as

$$\Lambda = \frac{\Lambda_0}{(1 + \tilde{u}_{rb0})^2}. \tag{22}$$

Like Λ_0, Λ is a ratio of electromagnetic to compression forces, but it is based on the radius at the deformed boundary, instead of the original boundary. The expansions given in Eqs. (19)–(20) are developed based on the steady stress distributions given in Eq. (10); however, they are approximately valid even in the transient case provided that $\partial\tilde{u}_{r0}/\partial\tilde{r} \approx \tilde{u}_{r0}/\tilde{r}$ when evaluated at the boundary. Application of these boundary conditions to the solutions given in Eq. (18) results in two linear homogeneous algebraic equations for the constants c_1 and c_2, having nontrivial solutions only when the determinant formed from these two equations is zero. The resulting dispersion relation, which may be found in the Appendix, defines the (ω, k) space where solutions exist for this eigenvalue problem.

Although solutions for $\tilde{u}_r$ and $\tilde{u}_z$ given in Eq. (18) are expressed in terms of modified Bessel functions, they are also applicable in situations where β is imaginary, after the use of identities relating ordinary to modified Bessel functions. Moreover, since solutions for $\tilde{u}_r$ and $\tilde{u}_z$ are necessarily real, the values for β and ω must either be purely real or imaginary. Thus, the displacements given in Eq. (18) define two families of solutions: modes which are unstable and grow in amplitude over time, and modes which are temporally oscillating and neither grow or decay in amplitude with time.

Limiting cases of the solutions given in Eq. (18) are also of interest. In particular, the long wavelength ($k \to 0$) limit yields

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

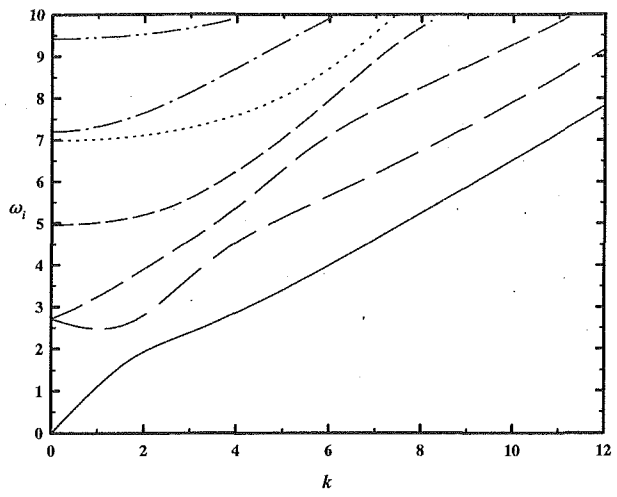


Fig. 3 Normalized frequency ω_i versus normalized wave number k for elastic disturbances in a cylinder with $\Phi = \frac{1}{2}$, $\Theta = \frac{5}{3}$ and $\Lambda = 0$

solutions for disturbances with vanishingly small axial dependences. In this situation the displacements reduce to

$$\tilde{u}_r = c_1 I_1\left(\frac{\omega\tilde{r}}{\sqrt{\Theta}}\right), \quad \tilde{u}_z = \left[\lim_{k\to 0}\frac{c_2}{k}\right]\frac{i\omega}{\sqrt{\Phi}} I_0\left(\frac{\omega\tilde{r}}{\sqrt{\Phi}}\right), \quad (23)$$

where the boundary conditions in Eqs. (19)–(20) are used to solve for the constants and determine the eigenvalues for ω. The dispersion relation that results may be found in the Appendix. In addition, the neutral stability ($\omega = 0$) limit determines solutions separating the unstable from the oscillating perturbation modes. The radial and axial displacements in this case are given as

$$\tilde{u}_r = c_1 I_1(k\tilde{r}) + c_2\tilde{r} I_0(k\tilde{r}),$$

$$\tilde{u}_z = c_1 i I_0(k\tilde{r}) + c_2 i\left[\tilde{r} I_1(k\tilde{r}) + \frac{2\Theta}{k(\Theta - \Phi)} I_0(k\tilde{r})\right], \quad (24)$$

where the boundary conditions in Eqs. (19)–(20) are again used to form the determinant condition necessary to solve for the eigenvalues. The dispersion relation resulting from this determinant is also listed in the Appendix.

5 Results and Discussion

Solutions to the eigenvalue problems formulated in Section 4 for the elastic response of the cylinder were generated for a wide range of input conditions. Solutions given in this paper are limited to $\Phi = \frac{1}{2}$ and $\Theta = \frac{5}{3}$, which are representative values for many metallic materials. Separate solutions were generated for real and imaginary values of the parameter ω, physically corresponding to the distending and oscillatory instabilities that can occur in the cylinder. In the discussion that follows, ω_r and ω_i are the real and imaginary moduli of ω; and are identified as the normalized growth rate and normalized frequency, respectively.

Shown in Fig. 3 is the normalized frequency ω_i versus the normalized wave number k when no electric current is present ($\Lambda = 0$). The dimensional frequency f and wavelength λ_w are related to ω_i and k as $f = \omega_i\sqrt{K/\rho}/2\pi r_{b0}$ and $\lambda_w = 2\pi r_{b0}/k$, respectively. As is evident in the figure, a large number of solutions are possible under these conditions. Solutions shown on a common line in the figure exhibited very similar deformation characteristics and thus belong to the same oscillatory deformation mode. In most cases, the frequency of each deformation mode tends to increase as the wavelength diminishes, except for the second mode, where the frequency increases with increasing wavelength when the wavelength is very long. The results shown here are identical to solutions by Onoe et al. (1962).

When a constant electric current is introduced, the response of the cylinder to instabilities changes dramatically. Shown in Fig. 4 is the normalized frequency versus wave number when $\Lambda = 4$. Note that this value of Λ corresponds to the perturbation of a transient solution to u_{r0}, since for $\Phi = \frac{1}{2}$ the maximum value for Λ that can be achieved is $\frac{14}{9}$, if u_{r0} is time independent. As such, the solutions shown here are strictly applicable for the case where $\partial\tilde{u}_{r0}/\partial\tilde{r} \approx \tilde{u}_{r0}/\tilde{r}$ at the boundary. However, this condition is approximately satisfied even in the transient case as long as accelerations at the boundary remain relatively small. In Fig. 4, solutions for perturbations corresponding to a common deformation mode are shown on the same line. As is evident from the figure, a large number of solutions are still possible when an electric current is added, but the locus of frequencies belonging to a common deformation mode tends to fluctuate as the wavelength shortens. Comparing the results in Fig. 4 to Fig. 3, it appears that the introduction of an axial current tends to decrease the frequency of disturbances at a given wavelength and mode. This effect is particularly prevalent for the lowest instability mode and at long wavelengths. However, certain solutions at the long wavelength limit appear to be unaffected by the current; particularly the third, fourth, sixth, and seventh instability modes shown on the two figures.

Purely elastic instabilities that grow in amplitude over time are also possible if the axial electric current is sufficient in magnitude. Shown in Fig. 5 is the normalized growth rate ω_r versus normalized wave number k for several different values of the parameter Λ. The normalized growth rate can be related to the dimensional growth rate γ as $\gamma = \omega_r\sqrt{K/\rho}/r_{b0}$. As is evident from the figure, only a single deformation mode exists for exponential instabilities when the electric current is fixed. Moreover, a critical or most unstable wavelength emerges where the growth rate of instabilities achieves a maximum for a given level of current. This critical wavelength tends to decrease with increasing current, except at very large axial currents, where it remains relatively constant. Furthermore, below a certain level of current, it appears that distending instabilities are not possible, since no eigenvalues for this problem were found for $\Lambda < 2.29$.

It is also of particular interest to determine the critical values of k for the conditions of neutral stability, where $\omega_r = 0$. This condition can be determined from the locus of abscissa intercepts from solutions shown in Fig. 5. Generation of this data

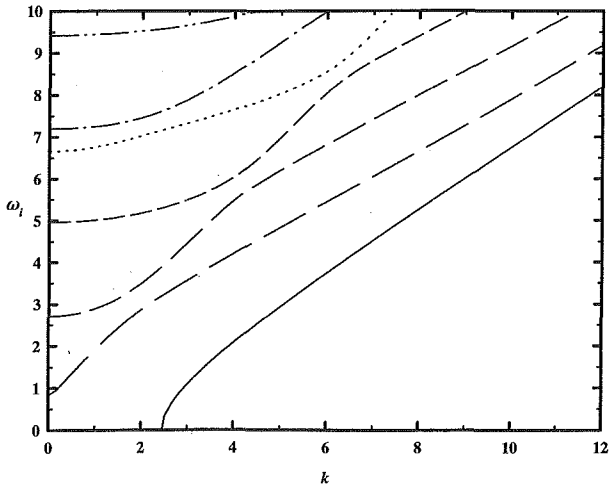


Fig. 4 Normalized frequency ω_i versus normalized wave number k for elastic disturbances in a cylinder with $\Phi = \frac{1}{2}$, $\Theta = \frac{5}{3}$ and $\Lambda = 4$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

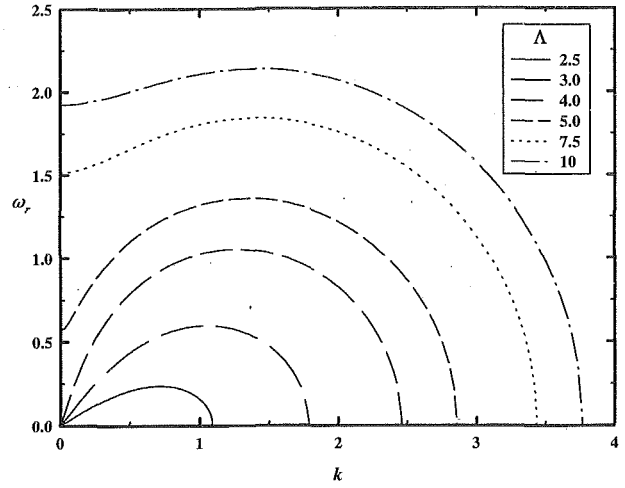


Fig. 5 Normalized growth rate ω_r versus normalized wave number k for elastic disturbances in a cylinder with $\Phi = \frac{1}{2}$ and $\Theta = \frac{5}{3}$

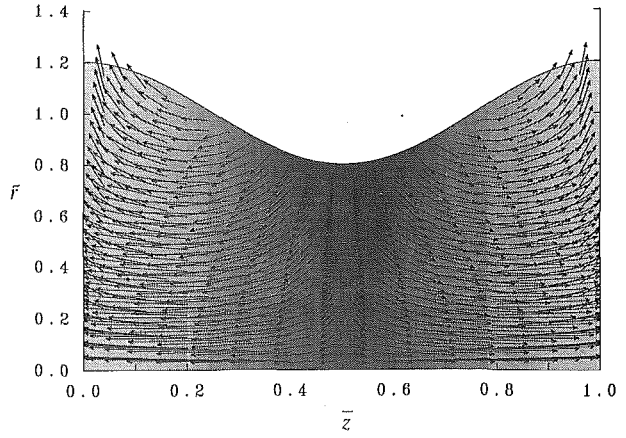


Fig. 7 Displacement and density distributions in a wavelength segment of the cylinder for the lowest oscillatory instability mode when $\Phi = \frac{1}{2}$, $\Theta = \frac{5}{3}$, $\Lambda = 0$ and $k = 2$ (displacement vectors are denoted with arrows, and density distributions with greyscale)

is of distinctive importance since it identifies the minimum current necessary to initiate an instability of a given wavelength. Shown in Fig. 6 is Λ versus k under the conditions of neutral stability. A minimum in the curve occurs at $k = 0$, where Λ is given by

$$\Lambda = \frac{12\Phi(\Theta - \Phi)}{3(2\Theta - \Phi) - 4(\Theta - \Phi)^2}. \qquad (25)$$

Below this value of Λ, purely elastic distending instabilities cannot be achieved.

The physical characteristics leading to instabilities in the elastic deformation of the cylinder can be more clearly recognized through an examination of the displacements associated with these solutions. Shown in Figs. 7 and 8 are first-order displacement vectors and density profiles for a wavelength segment of the cylinder when $k = 2$, $\Phi = 0.5$ and $\Lambda = 0$ for the first and second oscillatory instability modes, respectively. The density is calculated from a first-order expansion of mass conservation. The axial coordinate $\bar{z}$ is defined as $\bar{z} = z/\lambda_w$. Note that since the idealized displacements vanish when $\Lambda = 0$, the first-order displacements in this situation correspond to the actual deformation in the cylinder. Furthermore, since the zeroth-order density is constant, the first-order density is a measure of the stratification in density and pressure. Results in Fig. 7 indicate that the lowest instability mode is characterized by a displacement field consisting only of material flow out of regions of the cylinder where the cross sectional area is reduced. In addition, the density in these *necked* regions has a tendency to increase. The higher instability modes, on the other hand, exhibit multicellular behavior characterized by counter-rotating vortices in the interior of the cylinder, as is illustrated in Fig. 8. The density profiles are consistent with this deformation structure, indicating that the density increases in regions where displacement vectors have a tendency to converge. As was already stated, the frequency of these higher instability modes tends to be larger, as the eigenvalues for ω_i are 1.94 and 2.80 for the motions shown in Figs. 7 and 8, respectively.

First-order displacement and density eigenvectors were also computed for a number of other oscillatory modes and at different wavelengths, and for brevity will not be reproduced here. Instead, these results will be discussed in the context of the eigenvector solutions given in Figs. 7 and 8. When the value of Λ is held fixed and the wavelength is decreased, for example, the unicellular motion illustrated in Fig. 7 transitions to a displacement field consisting of multiple counter-rotating cells, as was shown in Fig. 8. Progressively shorter wavelength motions are characterized by transitions to an increasing number of

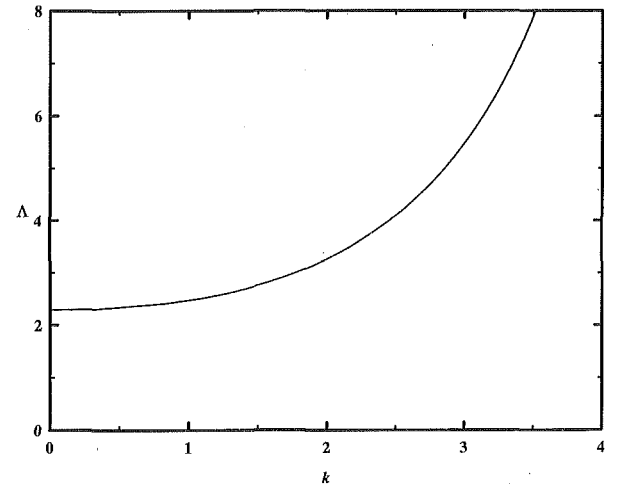


Fig. 6 Normalized current Λ versus normalized wave number k under neutral stability conditions for elastic disturbances in a cylinder with $\Phi = \frac{1}{2}$ and $\Theta = \frac{5}{3}$

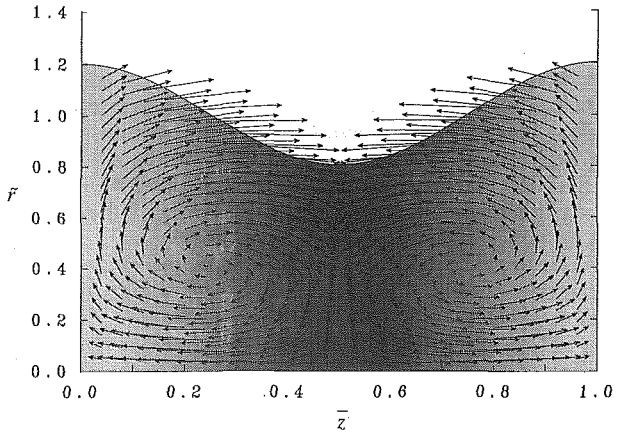


Fig. 8 Displacement and density distributions in a wavelength segment of the cylinder for the second oscillatory instability mode when $\Phi = \frac{1}{2}$, $\Theta = \frac{5}{3}$, $\Lambda = 0$ and $k = 2$ (displacement vectors are denoted with arrows, and density distributions with greyscale)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

counter-rotating cells across the radius. This behavior is exhibited consistently for all the oscillatory modes, but the value of k where the transitions occur changes. When the electric current is larger and the value of Λ increases, the displacement and density fields have characteristics very similar to those described above, but the transitions to multicellular motion occur at progressively shorter wavelengths. Thus, there appears to be a marked tendency for the electromagnetic field to suppress the multicellular displacement field.

Eigenvectors were also computed at the long wavelength ($k \to 0$) limit for oscillating perturbations. Solutions for the displacement and density fields separated into two very distinct motion types: in the first set of solutions, the axial displacement was zero and the density increased uniformly in necked regions of the cylinder, and in the second group, the radial displacement was zero and the density remained constant. Moreover, the electric current had no effect on disturbances belonging to this second group, as identical solutions were attained for all values of Λ.

As was shown in Fig. 5, purely elastic instabilities that grow in magnitude with time are also possible when the value of Λ is sufficiently large. Characteristics of the motion corresponding to these instabilities were determined to be very similar to the displacement and density fields shown in Figs. 7 and 8, so these results will be discussed in the context of the eigenvectors shown in these figures. For disturbances with long and moderate wavelengths, the motion in the cylinder is unicellular across the radius as was illustrated in Fig. 7. When the wavelength becomes short, however, the displacement field exhibits a multicellular behavior. When the electric current is increased, this transition occurs at progressively shorter wavelengths. For all values of Λ, the transition occurs at a value of k very close to the most unstable wavelength shown for the solutions in Fig. 5.

Examination of the deformation behavior in the cylinder, concurrently with the forces that generated the motion, suggests the foundation for the instabilities described in the preceding paragraphs. Apparently, a detailed consideration of the effects the axial stress has on the displacement and density distributions in necked regions of the cylinder is sufficient to account for the observed deformation behavior. The oscillatory perturbations, for instance, originate from the stress imbalance that results from both density increases and material displacement from necked regions of the cylinder. As material flows out of necked down regions, an axial tensile stress develops opposing the motion. In order to understand the source for this tensile stress, it is convenient to consider separately the mean and deviatoric components of the axial stress. A tensile stress, for instance, arises from the distortional component, since displacement directions in necked down regions necessarily require the deviatoric stress to be tensile. However, this deviatoric component must also be compressive in distended areas of the cylinder. Thus, an axial stress gradient develops, tending to direct the flow towards necked regions and restore the cylinder to its original configuration. Inertial forces cause an overshoot of the original undisturbed configuration and the motion oscillates. The frequency of this oscillation tends to decrease with wavelength, since the restoring stress gradient associated with shorter wavelength perturbations is larger. However, if the restoring axial stress gradient is sufficient in magnitude, it evidently causes material to flow back into necked regions of the cylinder, giving rise to the multicellular motion observed for the higher instability modes and at shorter wavelengths.

In order to determine the full effect of the axial stress on the stability of the cylinder, the mean component of the stress must also be considered. The effect of the mean component on the stability can be ascertained by considering the density change in necked regions. Since the density increase at necked regions contributes a compressive axial stress, it reduces in magnitude the distortional tensile stress that develops due to the flow. Thus, the principal effect of compressibility appears to be to lower the frequency of the oscillatory disturbances. The relative importance of these density changes is dependent on the magnitude of the bulk modulus; when the bulk modulus is large and density changes are small, more suppression of the oscillation frequency occurs.

In the absence of electromagnetic effects, material in the cylinder is always accelerated in the direction opposing its displacement and there is no tendency for instabilities to grow in amplitude. However, the addition of electric current significantly alters the stress distribution in the cylinder, as well as the resulting displacement field, and can cause perturbations to grow in magnitude under certain conditions. As material flows out of necked regions, for instance, a radial current develops along the surface of the cylinder, interacting with the azimuthal magnetic field in the vacuum to create an axial Lorentz force. The axial stress associated with this force is always compressive in necked regions, and thus tends to accelerate even more material out of the neck. If the current is sufficient in magnitude, evidently this acceleration tendency overcomes the natural restoring tendency arising from the distortional stress, and an unconditional instability results. This also suggests that the growth rate of the instability should depend on both the magnitude of the current and the wavelength of the disturbance, as the calculations have shown. The wavelength dependence, for instance, is the result of different wavelength dependencies for the natural restoring and Lorentz forces. The restoring stress gradient arising from the distortional stress is proportional to k^2 at very short wavelengths. The destabilizing stress gradient arising from Lorentz forces, on the other hand, is proportional to k, since for any given current the increase in the Lorentz force is proportional to the decrease in radius. Thus, the largest disparity between the two stress gradients occurs for perturbations with moderate values of k, which is consistent with the observed peaks in growth rate. For the oscillating perturbations, on the other hand, Lorentz forces tend to lower the frequency of perturbations, apparently since they lower the magnitude of the restoring axial stress gradient, for the reasons already indicated. This also suggests that frequency dampening should be most prevalent at moderate wavelengths and lower instability modes, where the disparity between restoring and Lorentz stress gradients is largest. The numerical results seem to be consistent with this observation. Apparently, electric currents have no effect on certain long wavelength instability modes, since they result only in density changes and no displacement of the radial boundary and, as a consequence, generate no axial gradients in the Lorentz force.

6 Conclusions

The results shown in the previous sections demonstrate that electromagnetic fields considerably change the motion of elastic cylinders. When the magnetic pressure exceeds a value of about three-fourths of the bulk modulus, a critical condition is exceeded and the radius of the cylinder is compressed down very rapidly. In addition, electromagnetic fields significantly alter the stability spectrum of elastic cylinders. Azimuthal magnetic fields have a tendency to restrain the frequency of perturbations that fluctuate in amplitude with time, and hasten the growth rate of instabilities that grow in magnitude over time. The magnetic pressure must be on the order of the bulk modulus in order for distending instabilities to occur. In most applications, however, the level of current required to yield the pressures needed to generate these effects is beyond the level of practical significance. However, as will be shown in the second part of this paper, distending instabilities can occur at much lower levels of current when the elastic-plastic behavior of the material is taken into account.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Acknowledgments

The technical work described in this paper was funded by the U.S. Army Research Office under contract No. DAAL03-92-K-0005, whose support is gratefully acknowledged.

References

Chace, W. G., and Moore, H. K., ed., 1968, *Exploding Wires,* Plenum, New York.

Chandrasekhar, S., 1961, *Hydrodynamic and Hydromagnetic Stability,* Clarendon Press, Oxford, U.K., pp. 515–574.

Chattopadhyay, S., and Moon, F. C., 1975, "Magnetoelastic Buckling and Vibration of a Rod Carrying Electric Current," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 809–914.

Graff, K. F., 1975, *Wave Motion in Elastic Solids,* Ohio State University Press, Columbus, OH.

Kruskal, M., and Schwartzschild, M., 1954, "Some Instabilities of a Completely Ionized Plasma," *Proceedings of the Royal Society of London,* Vol. A223, pp. 348–360.

Littlefield, D. L., and Powell, J. D., 1990, "The Effect of Electromagnetic Fields on the Stability of a Uniformly Elongating Plastic Jet," *Physics of Fluids A,* Vol. 2, No. 12, pp. 2240–2248.

Littlefield, D. L., 1991a, "Finite Conductivity Effects on the MHD Instabilities in Uniformly Elongating Plastic Jets," *Physics of Fluids A,* Vol. 3, No. 6, pp. 1666–1672.

Littlefield, D. L., 1991b, "Enhancement of Stability in Uniformly Elongating Plastic Jets with Electromagnetic Fields," *Physics of Fluids A,* Vol. 3, No. 12, pp. 2927–2935.

Littlefield, D. L., 1994, "Thermomechanical and Magnetohydrodynamic Stability of Elongating Plastic Jets," *Physics of Fluids,* Vol. 6, No. 8, pp. 2722–2729.

Littlefield, D. L., 1995, research in progress.

Littlefield, D. L., 1996, "Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part II: Plastic Response," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 742–749.

Malik, S. K., and Singh, M., 1994, "Stability of a Magnetic Fluid Jet," *International Journal of Engineering Science,* Vol. 29, No. 11, pp. 1493–1500.

Maugin, G. A., 1988, *Continuum Mechanics of Electromagnetic Solids,* Elsevier, New York.

Moon, F. C., 1984, *Magneto-solid Mechanics,* John Wiley and Sons, New York.

Nowacki, J. P., 1990, "Vibrating String in a Magnetic Field," *International Journal of Applied Electromagnetics in Materials,* Vol. 1, pp. 127–133.

Onoe, M., McNiven, H. D., and Mindlin, R. D., 1962, "Dispersion of Axially Symmetric Waves in Elastic Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, pp. 729–734.

Rayleigh, 1877, "The Theory of Sound," *London Mathematical Society Proceedings,* Vol. 17 (Dover Reprint, New York, 1945).

APPENDIX

Dispersion Relations for the Elastic Response

The solutions for the elastic response of the cylinder are given in Eq. (18). The arbitrary constants c_1 and c_2 are determined from substitution of the boundary conditions, given in Eqs. (19)–(20). Substitution of these boundary conditions results in two linear homogeneous algebraic equations, having nonzero solutions only when the determinant formed from these two equations is zero. This yields the dispersion relation

$$(\beta_2^2 + k^2)\left[\beta_2^2 + k^2 - \frac{k^2 Z}{2(\Theta - \Phi)}\right] I_0(\beta_1) I_1(\beta_2)$$

$$- 4\beta_1\beta_2 k^2\left[1 - \frac{Z}{4(\Theta - \Phi)}\right] I_0(\beta_2) I_1(\beta_1)$$

$$- 2\beta_1(\beta_2^2 - k^2)\left[1 + \frac{Z}{2\Phi}\right] I_1(\beta_1) I_1(\beta_2) = 0, \quad \text{(A1)}$$

where $\beta_1^2 = k^2 + \omega^2/\Theta$, $\beta_2^2 = k^2 + \omega^2/\Phi$, and $Z = 4(\Theta - \Phi)\Lambda/[4(\Theta - \Phi) + \Lambda]$. Solutions to this equation designate the (ω, k) values where solutions to the eigenvalue problem exist.

At the long wavelength limit, the dispersion relation can be simplified from the expression given in Eq. (A1). This equation can be expressed as

$$I_1\left(\frac{\omega}{\sqrt{\Phi}}\right)\left[\omega\sqrt{\Theta} I_0\left(\frac{\omega}{\sqrt{\Theta}}\right) - 2\left(\Phi + \frac{Z}{2}\right) I_1\left(\frac{\omega}{\sqrt{\Theta}}\right)\right] = 0. \quad \text{(A2)}$$

The I_1 multiplier contributes to nonzero roots for ω only when ω is purely imaginary, since $I_1(\beta) = -iJ_1(i\beta)$. Similarly, at the neutrally stable ($\omega = 0$) limit, the dispersion relation becomes

$$\left[(\Theta - \Phi)k^2 + \Theta - \frac{Z}{4}\left(k^2 - \frac{2\Theta}{\Phi}\right)\right] I_1^2(k)$$

$$- \frac{k\Theta Z}{4(\Theta - \Phi)} I_0(k) I_1(k)$$

$$- k^2\left(\Theta - \Phi - \frac{Z}{4}\right) I_0^2(k) = 0. \quad \text{(A3)}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part II: Plastic Response

D. L. Littlefield
Southwest Research Institute,
San Antonio, TX 78228
Mem. ASME

The analysis of elastic instabilities in metal cylinders when subjected to electromagnetic fields (Littlefield, 1996a) is extended in this work to include elastic-plastic flow. The cylinder is assumed to be infinitely long and perfectly conducting. The Prandtl-Reuss elastic-plastic material model is the assumed constitutive law, with the von Mises yield criterion employed to limit the effective stress. An axial electric current, assumed to be conducting along the surface of the cylinder, generates a confining pressure, causing plastic flow that is initially assumed to be uniform throughout the cross section. The propagation of small axisymmetric disturbances to this uniform motion is studied by applying linear perturbation theory. Solutions to these equations exhibit a wide range of instability modes, as was the case for the purely elastic results, and the frequency of the oscillating disturbances appears to be suppressed by electromagnetic effects. However, in contrast to the elastic result, no threshold magnetic field exists, and distending instabilities are possible for all levels of electric current. Physical mechanisms resulting in these distinctions are suggested.

1 Introduction

Electromagnetic fields can have a profound effect on the propagation of instabilities in cylinders. This phenomenon has been identified in virtually every material that conducts electricity, including magnetic fluids, semiconductors, plasmas, and metals. An understanding of this effect is important to many diverse industrial applications, such as the manufacture of fuses, microprocessor failures, magnetically confined fusion, and penetrator disruption. All of these applications may involve plastic flow of a metal, so in this paper, the elastic analysis from Part I is augmented to determine the effects of plastic flow on the stability of elastic-plastic cylinders, when exposed to large electromagnetic fields.

The study of stress wave propagation in metallic materials began many years ago with the classical treatment by Rayleigh (1877). The reader is referred to Graff's (1975) review for a detailed discussion of stress wave propagation in elastic materials. Studies of stress wave propagation in elastic-plastic materials began to surface after constitutive equations for elastic-plastic flow were established. The application usually considered in these analyses was dynamic impact problems, which is different from the problem considered here, where plastic flow is the result of an applied electromagnetic body force. Nevertheless, there are similarities between the two problems, since in both situations plastic flow is present and material inertia cannot be neglected. A comprehensive review of research in inelastic stress wave propagation for solids is given in Nowacki (1978).

Another application associated with research in disturbance propagation for cylindrical geometries has been the analysis of instabilities in stretching plastic jets, which has received considerable attention over the past decade. The propagation of small disturbances to the uniform stretching motion was usually considered in these analyses, so in most cases linearized versions of the governing equations were used. The original work in this area, reported by Chou and Carleone (1977), involved the use of a quasi-one-dimensional model to examine the stability of perfectly plastic jets when subjected to small perturbations. The analysis predicted that uniformly stretching metal jets were unstable with respect to all disturbances, and that very short wavelength perturbations possessed the largest growth rates. However, this conclusion was not consistent with the results of experiments, where disturbances of moderate wavelengths were observed to be the most unstable. Presumably, this disagreement motivated improvements to the original approach, one of which was implemented by Yarin (1994), where a more realistic constitutive model was employed. The improved model demonstrated how perturbations of finite wavelength could be the most unstable. Two-dimensional linear stability calculations were performed by Pack (1988), Romero (1989), and Littlefield and Powell (1990), who all predicted the emergence of a critical wavelength where perturbations tend to grow at the fastest rate. Moreover, the analysis by Littlefield and Powell suggested that the foundation for the emergence of this critical wavelength was the stabilizing stress enhancement that occurs for short wavelength perturbations.

It is well known that the application of externally generated body forces can alter the wave propagation characteristics in materials, irrespective of the source of this body force. For example, the application of magnetic fields can introduce instabilities in elastic structures that would otherwise be stable; the reader is referred to the texts by Moon (1984) and Maugin (1988) for a review of research in this area. It was shown in Part I of this paper that the application of magnetic fields to elastic cylinders alters the frequency dependence of oscillating perturbations, and introduces distending instabilities above a certain magnetic field threshold. However, if the magnetic field becomes too large, those results are no longer applicable since the material will yield and plastic flow will be initiated. Certainly the condition of plastic flow is expected to alter the conclusions reached for the elastic cylinder. Results that have already been established for plastic jets (Littlefield, 1990, 1991, 1994), for example, showed that the jets are unstable for all levels of azimuthal magnetic field. The purpose of this paper is to determine the effects of plastic flow on the stability of the cylinder. An infinitely long, perfectly conducting, and isother-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Dec. 19, 1994; final revision, Oct. 20, 1995. Associate Technical Editor: W. J. Drugan.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

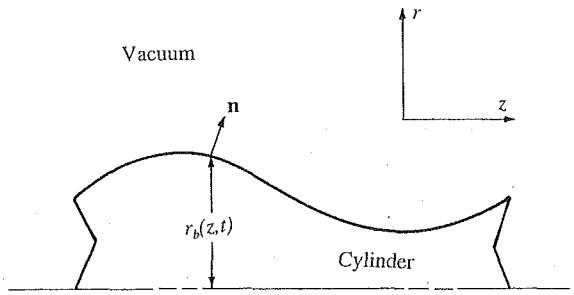


Fig. 1 Model for stability calculations

mal cylinder is considered. The assumed constitutive law is the Prandtl-Reuss formulation for an isotropic elastic-perfectly plastic material, and the elastic moduli and yield strength are further assumed to be constant. The elastic portion of the strain in the cylinder is assumed small and inertial effects are neglected; however, the geometric nonlinearity associated with the Lorentz force is incorporated by evaluating the magnetic stress at the deformed boundary, instead of the original boundary. The yield surface is assumed to conform to the von-Mises yield criterion. Adiabatic heating due to plastic work and resistive Joule heating are also neglected, so energy conservation is not considered in this analysis. An electric current is introduced to the cylinder, assumed to be sufficient in magnitude to cause the material to flow plastically. The governing equations admit a simple solution in this situation where the deformation in the cylinder remains in a state of uniaxial strain. The stability of this flow is then determined by considering the propagation of small, axisymmetric disturbances to the uniform motion. The applicable perturbation equations are solved to determine how the amplitude of these disturbances evolves over time.

The remainder of this paper is arranged as follows. The mathematical models for the cylinder and surrounding vacuum are presented in Section 2. In Section 3, these equations are solved under the idealized conditions of uniaxial strain. The equations governing the motion of small disturbances to the flow are derived and solved in Section 4. In Section 5, the solutions are presented for a number of situations, and physical mechanisms causing the observed behavior are suggested. Section 6 contains a brief conclusion, as well as the discussion of an application of these results.

2 Mathematical Model

Consider the motion of an axisymmetric cylinder as is depicted in Fig. 1. The radius at an arbitrary point on the surface of the cylinder is denoted as $r_b(z, t)$, and the outward unit normal vector at this point as $\mathbf{n}$. An axial electric current is assumed to be conducting along the surface of the cylinder. This current generates an azimuthal magnetic field in the surrounding vacuum, which interacts with the current to create a radial Lorentz force, directed towards the axis of the cylinder. The equations governing the motion in the cylinder may be expressed as

$$\rho \frac{\partial v_r}{\partial t} = -\frac{\partial p}{\partial r} + \frac{\partial \tau_{rr}}{\partial r} + \frac{2\tau_{rr} + \tau_{zz}}{r} + \frac{\partial \tau_{rz}}{\partial z} \tag{1}$$

and

$$\rho \frac{\partial v_z}{\partial t} = -\frac{\partial p}{\partial z} + \frac{\partial \tau_{rz}}{\partial r} + \frac{\tau_{rz}}{r} + \frac{\partial \tau_{zz}}{\partial z}, \tag{2}$$

where ρ is the density; v_r and v_z are the radial and axial velocities, respectively; p is the pressure; and τ_{ij} denotes a deviatoric component of the stress tensor. Equations (1) and (2) are the equations for conservation of radial, and axial momentum, respectively; assuming negligible inertial effects. The electromagnetic body force does not appear in the momentum equations, which is a result of the assumption on the conductivity; since the electric current was initially assumed to be concentrated on the surface, it must remain on the surface if the conductivity is assumed infinite. The pressure and stress tensor appearing in Eqs. (1)–(2) are determined by applying the Prandtl-Reuss constitutive model, and are given as

$$\left.\begin{aligned} \frac{\partial v_r}{\partial r} &= \frac{1}{2G}\frac{\partial \tau_{rr}}{\partial t} - \frac{1}{3K}\frac{\partial p}{\partial t} + \frac{\tau_{rr}}{2\lambda} \\ \frac{\partial v_z}{\partial z} &= \frac{1}{2G}\frac{\partial \tau_{zz}}{\partial t} - \frac{1}{3K}\frac{\partial p}{\partial t} + \frac{\tau_{zz}}{2\lambda} \\ \frac{\partial v_r}{\partial z} + \frac{\partial v_z}{\partial r} &= \frac{1}{G}\frac{\partial \tau_{rz}}{\partial t} + \frac{\tau_{rz}}{\lambda} \\ \frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{\partial v_z}{\partial z} &= -\frac{1}{K}\frac{\partial p}{\partial t} \end{aligned}\right\}, \tag{3}$$

where λ is a positive scalar function describing the material flow resistance to applied shear. Note that since the problem under consideration can be described as *self-contained* plastic deformation, the pressure appearing in Eq. (3) may be quite large. Thus, in writing Eq. (3), the elastic portions of the strain must also be assumed to be small. The stress tensor in Eq. (3) must satisfy the von Mises yield condition, given by

$$\tau_{rr}^2 + \tau_{rr}\tau_{zz} + \tau_{zz}^2 + \tau_{rz}^2 = \frac{Y^2}{3}, \tag{4}$$

and Y is the constant yield strength. Here is it assumed that the electric current is sufficient in magnitude to sustain plastic flow; the current necessary to satisfy this condition will be formulated in the next section. In the vacuum surrounding the cylinder, the magnetic field and must satisfy the magnetic diffusion equation given as

$$\frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial}{\partial r}(rB_\theta)\right] + \frac{\partial^2 B_\theta}{\partial z^2} = 0, \tag{5}$$

where B_θ is the azimuthal component of the magnetic field. Equation (5) is formed in a straightforward manner from Maxwell's equations under the assumptions of cylindrical symmetry and vanishing azimuthal current density and electric field.

Application of Eqs. (1)–(5) implies that certain boundary conditions must be satisfied at the interface between the cylinder and vacuum. In particular, mass and momentum must be conserved and Maxwell's equations must be satisfied. The required boundary conditions can be systematically derived by integrating the general forms of the governing equations, applicable in both the cylinder and the surrounding vacuum, across the boundary, presumed to be of thickness Δ, followed by a limit process where $\Delta \to 0$ (e.g., see Kruskal and Schwartzchild, 1954). This insures that the conservation equations will still be satisfied at the boundary. The conditions that result from this process are given in Part I of this paper (Littlefield, 1996a), and for convenience are repeated here:

$$\frac{\partial r_b}{\partial t} = v_r - \frac{\partial r_b}{\partial z} v_z, \tag{6}$$

$$-p + \tau_{rr} + \frac{B_\theta^2}{2\mu} - \frac{\partial r_b}{\partial z}\tau_{rz} = 0, \tag{7}$$

$$\frac{\partial r_b}{\partial z}\left(-p + \tau_{zz} + \frac{B_\theta^2}{2\mu}\right) - \tau_{rz} = 0, \tag{8}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and

$$B_\theta = \frac{\mu I}{2\pi r}, \tag{9}$$

where I is the constant electric current, μ is the magnetic permeability, and these conditions are all to be evaluated at $r = r_b$. The pressure and stress terms that appear in these equations are evaluated using Eq. (3). Together with these boundary conditions, Eqs. (1)–(5) represent a closed system of nine coupled partial differential equations that can be solved for the nine unknowns in the cylinder and vacuum: v_r, v_z, p, τ_{rr}, τ_{rz}, τ_{zz}, λ, r_b, and B_θ. In their most general form, these nonlinear equations are formidable and quite difficult so solve. However, a reduced set of linear equations may be solved when considering small disturbances to the simplified idealized motion of the cylinder.

3 Idealized Motion

The deformation in the cylinder studied in this analysis is the result of an imposed radially varying magnetic pressure, that develops due to the interaction of the axial electric current and azimuthal magnetic field. The idealized deformation is assumed to be in a yielded uniaxial strain state. The constraint on the current, necessary to insure plastic flow in the cylinder, can be determined by substituting solutions for the elastic stresses in the cylinder (Littlefield, 1996a) for steady uniaxial strain into the yield condition, given in Eq. (4). When expressed in dimensionless form, this condition is given by

$$\Lambda_0 \geq 2\Psi\left(1 + \frac{\Phi}{3}\right)\left(1 - \frac{\Psi}{2}\right)^2, \tag{10}$$

where Λ_0, Φ, and Ψ are dimensionless parameters defined as

$$\Lambda_0 = \frac{\mu I^2}{4\pi^2 r_{b0}^2 K}, \quad \Phi = \frac{G}{K}, \quad \Psi = \frac{Y}{G}. \tag{11}$$

In this relation, K and G are the elastic bulk and shear moduli, respectively; and the 0 subscript is used to denote an idealized quantity. Physically, Λ_0 represents a ratio of electromagnetic to compressive forces, and Φ a ratio of shear to compressive forces. Ψ is a ratio of plastic forces to elastic shear forces, the Deborah number, and is typically very small for the problem at hand. Solutions to Eqs. (1) and (3)–(4), together with the boundary conditions in Eqs. (6)–(7) and (9), determine the motion in the cylinder. They require the radial and axial velocities in the cylinder to be zero, thus the radial and axial displacements u_r and u_z are time independent. The deviatoric stress and pressure under these conditions are given by

$$\tilde{\tau}_{rr0} = -\Phi\Psi/3; \quad \tilde{\tau}_{zz0} = 2\Phi\Psi/3; \quad \tilde{\tau}_{rz0} = 0; \quad \tilde{p} = -2\tilde{u}_{rb0}; \tag{12}$$

where $\tilde{\tau} = \tau/K$ and $\tilde{p} = p/K$. $\tilde{u}_{rb0}$ is the value of the normalized radial displacement $\tilde{u}_{r0}$ evaluated at the boundary $r = r_{b0}$, where $\tilde{u}_{r0} = u_{r0}/r_{b0}$. With use of Eq. (12), Eq. (7) requires that $\tilde{u}_{rb0}$ satisfy

$$\tilde{u}_{rb0}^3 + 2\left(1 - \frac{\Psi\Phi}{12}\right)\tilde{u}_{rb0}^2 + \left(1 - \frac{\Psi\Phi}{3}\right)\tilde{u}_{rb0} + \frac{\Lambda_0}{4} - \frac{\Psi\Phi}{6} = 0. \tag{13}$$

Equation (13) yields three solutions for the radial displacement, but only one of these is physically meaningful, given by

$$\tilde{u}_{r0} = \tilde{u}_{rb0}\frac{r}{r_{b0}} = -\frac{2}{3}\left\{1 - \frac{\delta}{4} - \left(1 + \frac{\delta}{2}\right) \times \cos\left[\frac{1}{3}\tan^{-1}\left(\frac{36}{\sqrt{6}}\frac{\sqrt{\alpha(\delta+2)^3 - 54\alpha^2}}{(\delta+2)^3 - 108\alpha}\right)\right]\right\}\frac{r}{r_{b0}}, \tag{14}$$

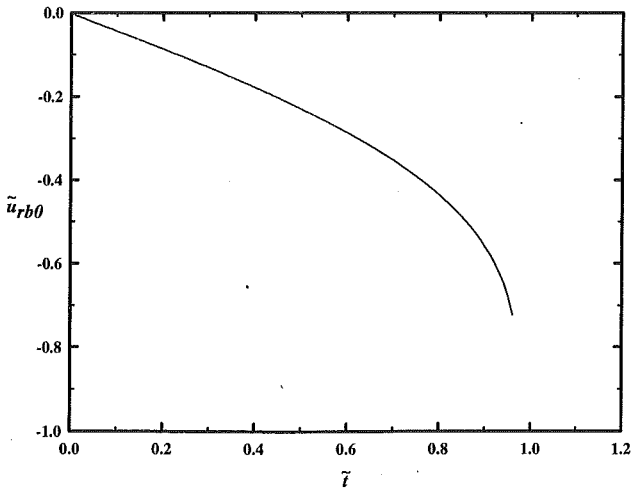


Fig. 2 Normalized displacement at the boundary $\tilde{u}_{rb0}$ versus normalized time $\tilde{t}$ for an elastic-plastic cylinder with $\Phi = \frac{1}{2}$, $\Psi = 0$ and $\Lambda = 0.8$

where $\alpha = \Lambda_0/4$ and $\delta = \Psi\Phi/3$. This solution possesses many of the same characteristics as the solution for $\tilde{u}_{r0}$ given in Eq. (13) from Part I of this paper. For example, real solutions to Eq. (14) exist for $\alpha < \alpha_c = (\delta + 2)^3/54$, and above this limit, there are no steady-state solutions for $\tilde{u}_{r0}$. The transient solutions found for $\alpha > \alpha_c$ exhibit very rapid decreases in displacement when $\tilde{u}_{r0}$ exceeds a critical level corresponding to its steady-state value when $\alpha = \alpha_c$. Critical values for $\tilde{u}_{rb0}$ and Λ_0 are derived from Eq. (14) by setting $\alpha = \alpha_c$ and are given by

$$\tilde{u}_{rb0c} = -\frac{1}{3} + \frac{\Phi\Psi}{9}; \quad \Lambda_{0c} = \frac{16}{27}\left(1 + \frac{\Phi\Psi}{6}\right)^3; \tag{15}$$

where the c subscript denotes a critical value.

If the current is large enough so that $\Lambda_0 > \Lambda_{0c}$, then transient solutions for $\tilde{u}_{r0}$ must be considered. Shown in Fig. 2 is a plot of $\tilde{u}_{rb0}$ as a function of normalized time $\tilde{t}$, where $\tilde{t} = t\sqrt{K/\rho}/r_{b0}$. The values for Λ_0, Φ, and Ψ were 0.8, 0.5, and 0, respectively. The rapid decrease in boundary displacement when $\tilde{u}_{rb0}$ drops below its critical value of $\frac{1}{3}$ is evident. As was already identified from Part I of the paper, the rapid decline is caused by the intensification in magnetic pressure as $\tilde{u}_{rb0}$ approaches its critical value. At that point, the magnetic pressure increasingly overmatches the restoring stress, and the radius of the cylinder races down very quickly. Admittedly, the solution shown in Fig. 2 is not strictly valid since the elastic strains in the cylinder are not small; however, preliminary calculations performed for finite strains in elastic cylinders (Littlefield, 1995) have shown that the essential behavior illustrated in Fig. 2 is retained even when finite strains are taken into account.

In the vacuum surrounding the cylinder, only the magnetic field needs to be determined, and this field must be independent of z in order to be consistent with the uniaxial strain conditions assumed. Therefore, integration of Eq. (5) subject to the boundary condition given in Eq. (9), along with the requirement that $B_{\theta 0}$ remain finite as $r \to \infty$, yields

$$B_{\theta 0} = \frac{\mu I}{2\pi r} \tag{16}$$

for the azimuthal magnetic field.

4 First-Order Equations and Solutions

Perturbations to the solutions developed in the previous section are now considered. The equations that govern the motion

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of these disturbances are developed using linear perturbation theory. These solutions remain valid as long as the amplitude of the disturbances remains small.

If the electric current is sufficient in magnitude, the yield criterion specified in Eq. (10) is satisfied and plastic flow initiates. As was previously stated, the idealized steady motion under this condition is characterized by constant radial and azimuthal strains that increase with the level of current. After they are formed, the equations governing perturbations to this motion turn out to have coefficients independent of both z and t, so the axial and temporal dependences are both separable and exponential. Using the first-order variables defined as

$$\left.\begin{aligned}
\lambda_1 &= \tilde{\lambda}\,\frac{\lambda_0^2}{Yr_{b0}}\sqrt{K/\rho}\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
v_{r1} &= \tilde{v}_r\sqrt{K/\rho}\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
v_{z1} &= \tilde{v}_z\sqrt{K/\rho}\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
p_1 &= \tilde{p}K\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
\tau_{rr1} &= \tilde{\tau}_{rr}K\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
\tau_{zz1} &= \tilde{\tau}_{zz}K\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})\\
\tau_{rz1} &= \tilde{\tau}_{rz}K\,\exp(ik\tilde{z})\,\exp(\omega\tilde{t})
\end{aligned}\right\}, \tag{17}$$

where $\tilde{z} = z/r_{b0}$, k is the axial wave number, and ω is the normalized growth rate, then Eqs. (1)–(4) yield

$$\omega\tilde{v}_r = -\frac{\partial\tilde{p}}{\partial\tilde{r}} + \frac{\partial\tilde{\tau}_{rr}}{\partial\tilde{r}} + \frac{2\tilde{\tau}_{rr} + \tilde{\tau}_{zz}}{\tilde{r}} + ik\tilde{\tau}_{rz}, \tag{18}$$

and

$$\omega\tilde{v}_z = -ik\tilde{p} + \frac{\partial\tilde{\tau}_{rz}}{\partial\tilde{r}} + \frac{\tilde{\tau}_{rz}}{\tilde{r}} + ik\tilde{\tau}_{zz}, \tag{19}$$

where $\tilde{r} = r/r_{b0}$, and the pressure and deviatoric stress satisfy the relations

$$\left.\begin{aligned}
&\frac{\partial\tilde{v}_r}{\partial\tilde{r}} = \omega\left(\frac{\tilde{\tau}_{rr}}{2\Phi} - \frac{\tilde{p}}{3}\right) + \frac{\tilde{\lambda}}{6}\\
&ik\tilde{v}_z = \omega\left(\frac{\tilde{\tau}_{zz}}{2\Phi} - \frac{\tilde{p}}{3}\right) - \frac{\tilde{\lambda}}{3}\\
&ik\tilde{v}_r + \frac{\partial\tilde{v}_z}{\partial\tilde{r}} = \frac{\omega\tilde{\tau}_{rz}}{\Phi}\\
&\frac{\partial\tilde{v}_r}{\partial\tilde{r}} + \frac{\tilde{v}_r}{\tilde{r}} + ik\tilde{v}_z = -\omega\tilde{p}\\
&\tilde{\tau}_{zz} = 0
\end{aligned}\right\}. \tag{20}$$

Equation (20) is derived based on perturbations from the steady solutions for the stress given in Eq. (12). As such, the stresses given in Eq. (20) assume that no elastic unloading occurs as a result of the applied perturbations. Normally this is not a significant limitation because, as can be seen from Eq. (12), a small change in displacement does not result in elastic unloading, but instead changes the pressure in the cylinder. Equations (18)–(20) may be combined into a single equation for the radial velocity, given by

$$\left\{(\omega^2 + \Phi k^2)(\omega^2 + k^2) - [\Phi(\omega^2 - k^2) + \Omega\omega^2]\left(\frac{\partial^2}{\partial\tilde{r}^2} + \frac{1}{\tilde{r}}\frac{\partial}{\partial\tilde{r}} - \frac{1}{\tilde{r}^2}\right) + \Phi\Omega\left(\frac{\partial^2}{\partial\tilde{r}^2} + \frac{1}{\tilde{r}}\frac{\partial}{\partial\tilde{r}} - \frac{1}{\tilde{r}^2}\right)^2\right\}\tilde{v}_r = 0, \tag{21}$$

where $\Omega = 1 + \Phi$. Physically, Ω and Φ represent normalized longitudinal and transverse wave speeds of plastic waves at the long wavelength ($k \to 0$) limit, but these wave speeds are not applicable to intermediate or short wavelength disturbances. This is in contrast to the elastic result, where the longitudinal and transverse wave speeds are fixed and independent of wavelength. Indeed, for certain wavelength perturbations, no real wave speeds for plastic disturbances exist, as the form of Eq. (21) can no longer be identified as hyperbolic. In all cases where real plastic wave speeds were established, however, the wave speeds were also slower than or equal to their corresponding elastic counterparts. These results have also been suggested in previous analyses (Bleich and Nelson, 1966), where the propagation velocities for plastic waves were determined to be highly dependent on the geometry and loading condition, and were always slower than or equal to the corresponding elastic wave speeds.

The solution to Eq. (21) is dependent on the roots of a characteristic polynomial for β, given by

$$(\omega^2 + \Phi k^2)(\omega^2 + k^2) + [\Phi(\omega^2 - k^2) + \Omega\omega^2]\beta^2 + \Phi\Omega\beta^4 = 0. \tag{22}$$

When β^2 is real, the solutions for $\tilde{v}_r$ and $\tilde{v}_z$ are given by

$$\tilde{v}_r = c_1 J_1(\beta_1\tilde{r}) + c_2 J_1(\beta_2\tilde{r}),$$

$$\tilde{v}_z = c_1\frac{i\beta_1}{k}\left[1 + \frac{\omega^2 + \Phi k^2}{\Omega\beta_1^2}\right]J_0(\beta_1\tilde{r}) + c_2\frac{i\beta_2}{k}\left[1 + \frac{\omega^2 + \Phi k^2}{\Omega\beta_2^2}\right]J_0(\beta_2\tilde{r}), \tag{23}$$

where β_1 and β_2 are the two positive roots for β, and c_1 and c_2 are arbitrary constants. These solutions are also applicable when either β_1 or β_2 is purely imaginary, with the use of Bessel function identities. When β^2 is imaginary or complex, however, the solutions are

$$\tilde{v}_r = c_1 RJ_1(\beta\tilde{r}) + c_2 IJ_1(\beta\tilde{r}),$$

$$\tilde{v}_z = \frac{ic_1}{k}\left\{\beta_r\left[1 + \frac{\omega^2 + \Phi k^2}{\Omega\beta_m^2}\right]RJ_0(\beta\tilde{r}) - \beta_i\left[1 - \frac{\omega^2 + \Phi k^2}{\Omega\beta_m^2}\right]IJ_0(\beta\tilde{r})\right\} + \frac{ic_2}{k}\left\{\beta_i\left[1 - \frac{\omega^2 + \Phi k^2}{\Omega\beta_m^2}\right]RJ_0(\beta\tilde{r}) + \beta_r\left[1 + \frac{\omega^2 + \Phi k^2}{\Omega\beta_m^2}\right]IJ_0(\beta\tilde{r})\right\}, \tag{24}$$

where β_r and β_i are the magnitude of the real and imaginary portions of β, respectively; and $\beta_m^2 = \beta_r^2 + \beta_i^2$. The functions RJ and IJ are the real and imaginary portions of the ordinary Bessel function of a complex argument. Details of the solutions for $\tilde{v}_r$ and $\tilde{v}_z$ in this situation are nontrivial, and are the subject of an accompanying paper (Littlefield, 1996b). The constants c_1 and c_2 appearing in Eqs. (23)–(24) are determined through application of the appropriate boundary conditions. These conditions are developed from first-order expansions of Eqs. (6)–(9), with use of the constitutive relations given in Eq. (20), and are given as

$$\frac{1}{\tilde{r}}\frac{\partial}{\partial\tilde{r}}(\tilde{r}\tilde{v}_r) + ik\tilde{v}_z + \Phi\left(\frac{\partial\tilde{v}_r}{\partial\tilde{r}} - \frac{\tilde{v}_r}{\tilde{r}}\right) - Z\tilde{v}_r = 0 \tag{25}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and

$$ik(1 - \Psi)\frac{Z}{\Lambda}\tilde{v}_r + \frac{\partial \tilde{v}_z}{\partial \tilde{r}} = 0, \tag{26}$$

where

$$Z = \frac{4\Lambda}{4 + \Lambda - 2\Phi\Psi/3}, \tag{27}$$

and these conditions are evaluated at $\tilde{r} = 1$. The parameter Λ appearing in these equations is related to Λ_0 as

$$\Lambda = \frac{\Lambda_0}{(1 + \tilde{u}_{rb0})^2}, \tag{28}$$

where Λ is a ratio of electromagnetic to compressive forces evaluated at the *deformed* boundary, rather than the original boundary. Note that the parameter Ψ, which is a dimensionless measure of the yield strength, appears only in the boundary conditions. Furthermore, since for most metals $\Psi \ll 1$ the magnitude of the yield strength has virtually no effect on the deformation of the cylinder; it only designates the level of electric current needed to initiate plastic flow. However, the condition of plastic flow substantially alters solutions for the deformation in the cylinder, when compared to the purely elastic solutions, as will become evident in the next sections.

The boundary conditions given in Eqs. (25)–(26) are used to determine the constants c_1 and c_2 appearing in Eqs. (23)–(24). This yields two linear homogeneous algebraic equations, having nonzero solutions only when the determinant formed from the two equations is zero. The resulting dispersion relation is given in the Appendix. Solutions to this transcendental equation defines the (ω, k) space where solutions to this eigenvalue problem exist.

Since solutions for v_{r1} and v_{z1} cannot be complex, restrictions must be placed on the values for ω and k; in particular, k must always be real and ω must be either real or imaginary. Consequently, solutions to the eigenvalue problem for v_{r1} and v_{z1} can be grouped into two categories: solutions where ω is real, having amplitudes that grow exponentially over time, and solutions where ω is imaginary, possessing amplitudes that oscillate with a normalized frequency ω/i, and neither grow nor decay over time.

Limiting cases of the solutions given in Eqs. (23)–(24) are also of interest. The long wavelength limit yields solutions where both the idealized and perturbed flow fields remain in a condition of uniaxial strain. When $k = 0$, the solutions for the first-order velocities are given as

$$\tilde{v}_r = c_1 I_1\left(\frac{\omega\tilde{r}}{\sqrt{\Omega}}\right),$$

$$\tilde{v}_z = \left[\lim_{k\to 0}\frac{c_2}{k}\right]\frac{i\omega}{\sqrt{\Phi}}\left(1 + \frac{\Phi}{\Omega}\right)I_0\left(\frac{\omega\tilde{r}}{\sqrt{\Phi}}\right). \tag{29}$$

The boundary conditions in Eqs. (25)–(26) are applied to establish the dispersion relation. Similarly, when $\omega = 0$ the conditions for incipient instability are determined. In this situation the solutions given in Eqs. (23)–(24) reduce to

$$\tilde{v}_r = c_1 R J_1(\beta\tilde{r}) + c_2 I J_1(\beta\tilde{r}),$$

$$\tilde{v}_z = \frac{ic_1}{k}\left\{\beta_r\left[1 + \frac{\Phi}{\sqrt{\Omega}}\right]RJ_0(\beta\tilde{r}) - \beta_i\left[1 - \frac{\Phi}{\sqrt{\Omega}}\right]IJ_0(\beta\tilde{r})\right\}$$

$$+ \frac{ic_2}{k}\left\{\beta_i\left[1 - \frac{\Phi}{\sqrt{\Omega}}\right]RJ_0(\beta\tilde{r})\right.$$

$$\left. + \beta_r\left[1 + \frac{\Phi}{\sqrt{\Omega}}\right]IJ_0(\beta\tilde{r})\right\}, \tag{30}$$

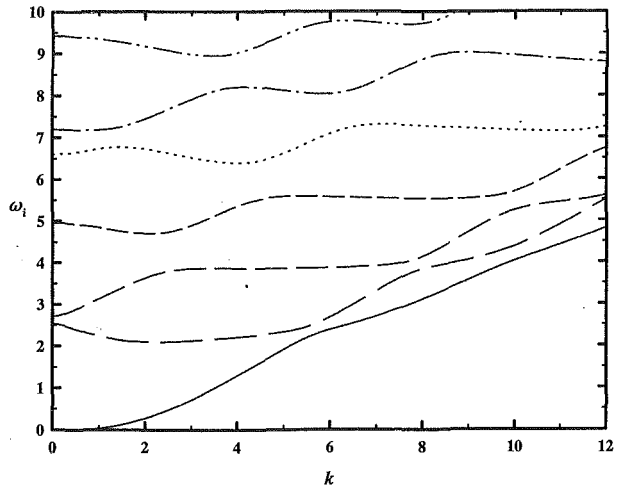


Fig. 3 Normalized frequency ω_i versus normalized wave number k for plastic disturbances in a cylinder with $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$, $\Psi = 0$ and $\Lambda = 0$

where $\beta_r^2 = (1 + 2\sqrt{\Omega})k^2/4\Omega$ and $\beta_i^2 = (2\sqrt{\Omega} - 1)k^2/4\Omega$. Application of the boundary conditions given in Eqs. (25)–(26) yields the dispersion relation needed to determine the eigenvalues for k and Λ, and is listed in the Appendix.

5 Results and Discussion

Solutions to the eigenvalue problems formulated in Section 4 for the plastic response of the cylinder were generated for a wide range of input conditions. Solutions given in this paper are limited to $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$ and $\Psi = 0$, which are representative values for many metallic materials. Thus the yield strength is not explicitly used as an input parameter in the analysis, but nevertheless the condition of plastic flow considerably alters the spectrum of solutions to the given equations. Separate solutions were generated for real and imaginary values of the parameter ω, physically corresponding to the distending and oscillatory instabilities that can occur in the cylinder. In the discussion that follows, ω_r and ω_i are the real and imaginary moduli of ω; and are identified as the normalized growth rate and normalized frequency, respectively.

Figure 3 depicts the solutions for oscillatory disturbances when $\Lambda = 0$, shown as normalized frequency versus normalized wave number. The normalized frequency and wave number are related to the dimensional frequency f and wavelength λ_w as $f = \omega_i\sqrt{K/\rho}/2\pi r_{b0}$ and $\lambda_w = 2\pi r_{b0}/k$, respectively. The condition $\Lambda = 0$ physically corresponds to a situation where the yield strength and current are both vanishingly small, but the current is still sufficient in magnitude to satisfy the plastic flow condition given in Eq. (10). Comparing the solutions shown in this figure to the results for elastic disturbances (see Littlefield, 1996a, Fig. 3) it is apparent that many features are common to both situations. A large number of oscillating instability modes are still possible when the cylinder is flowing plastically. Solutions can also be grouped into similar deformation modes, which are shown along common lines in the figure. However, the values for ω_i at a fixed value of k shown in Fig. 3 appear to be considerably lower than the corresponding elastic values, particularly at shorter wavelengths. Moreover, the frequency dependence for oscillating modes tends to fluctuate with wavelength, whereas ω_i predominantly increased with k for elastic disturbances.

Changes in frequency dependence of the oscillatory instability modes occurs when the current is increased. Shown in Fig. 4 is the normalized frequency versus normalized wave number when $\Lambda = 4$. Note that this value of Λ corresponds to a transient solution for the idealized motion, since when $\Psi = 0$ the maxi-

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

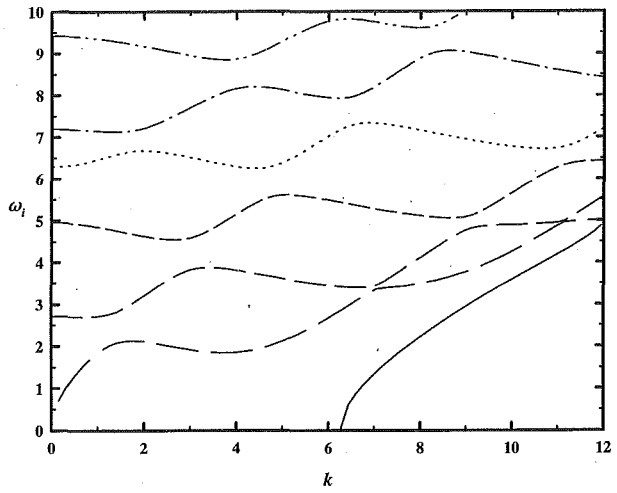


Fig. 4 Normalized frequency ω_i versus normalized wave number k for plastic disturbances in a cylinder with $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$, $\Psi = 0$ and $\Lambda = 4$

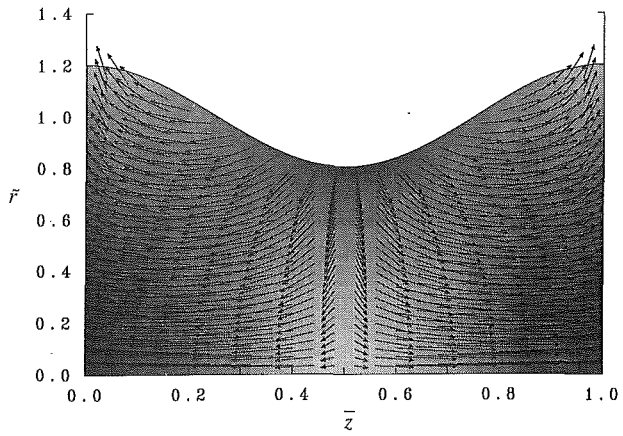


Fig. 6 Velocity and density distributions in a wavelength segment of the cylinder for the lowest oscillatory instability mode when $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$, $\Psi = 0$, $\Lambda = 0$ and $k = 2$ (velocity vectors are denoted with arrows, and density distributions with greyscale)

mum value that can be attained for Λ is $\frac{4}{3}$ under steady conditions. Thus, these solutions are not strictly applicable since were derived based on perturbations from the steady-state stresses given in Eq. (12). However, when the acceleration at the boundary is relatively small, the stress solution in Eq. (12) is approximately correct, as are the perturbation solutions. The results shown in the figure suggest that larger currents tend to diminish the frequency of disturbances at a given mode and wavelength. Apparently, the frequencies of lower modes are more strongly subdued by electromagnetic effects, and within a given mode more suppression seems to occur when ω_i is increasing with k. Although similar observations were also noted for solutions to the elastic response, the magnitude of the suppression appears to be more pronounced for plastic flow, particularly at the lowest instability mode.

Distending instabilities are also possible when plastic flow is present. Shown in Fig. 5 is the normalized growth rate ω_r versus normalized wave number k for various values of the parameter Λ. The normalized growth rate is related to the dimensional growth rate γ as $\gamma = \omega_r\sqrt{K/\rho}/r_{b0}$. In general, the qualitative behavior of these results seems to compare favorably with the elastic solutions (see Littlefield, 1996a, Fig. 5), but magnitudes of growth rates tend to be larger at all values of Λ. Moreover, instability occurs at all levels of electric current, whereas for the elastic response instability occurred only above a threshold value for Λ. At the long wavelength limit below $\Lambda = 4$, no nonzero values for ω_r existed. Evidently, the most unstable perturbations for the plastic response tend to monotonically diminish in wavelength as the electric current is increased.

Eigenvectors for the velocity and density distributions in the cylinder associated with the solutions shown in Figs. 3–5 were also computed. Shown in Figs. 6 and 7 are velocity and density distributions of a wavelength segment of the cylinder for the first and second oscillatory modes when $\Lambda = 0$ and $k = 2$. The density is calculated from a first-order expansion of mass conservation. The normalized axial coordinate $\bar{z}$ is defined as $\bar{z} = z/\lambda_w$. The results in Fig. 6 indicate that the lowest oscillatory mode is characterized by unicellular motion, where material flows out of necked down regions of the cylinder and into distended areas. The density in these necked regions also tends to increase. However, the flow characteristics of the higher instability modes are dramatically different from this behavior. In Fig. 7, where the second mode is shown, the presence of counter rotating vortices in the interior of the cylinder is evident. These vortices cause material to return to necked regions close to the radial boundary of the cylinder. The density is apparently

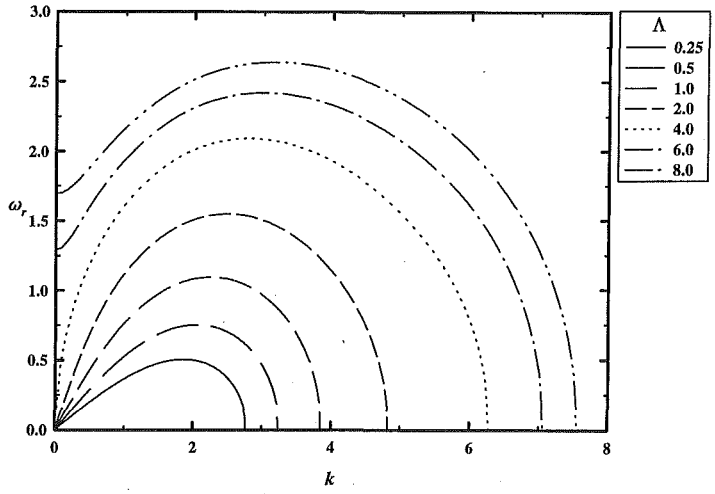


Fig. 5 Normalized growth rate ω_r versus normalized wave number k for plastic disturbances in a cylinder with $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$ and $\Psi = 0$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

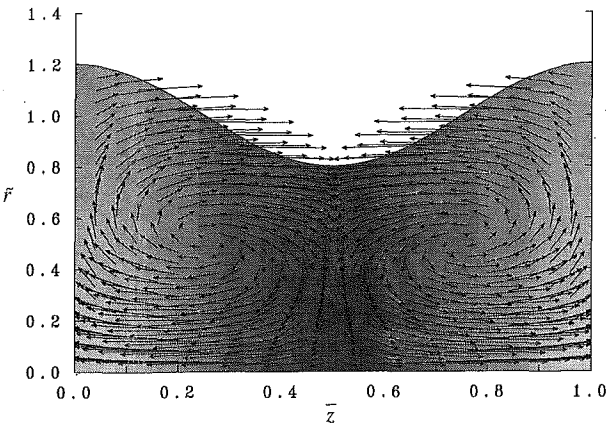


Fig. 7 Velocity and density distributions in a wavelength segment of the cylinder for the second oscillatory instability mode when $\Phi = \frac{1}{2}$, $\Omega = \frac{3}{2}$, $\Psi = 0$, $\Lambda = 0$ and $k = 2$ (velocity vectors are denoted with arrows, and density distributions with greyscale)

consistent with flow behavior, tending to increase in regions where the flow converges.

Eigenvectors for many values for the input parameters other than those shown in Figs. 6 and 7 were also computed, and for brevity will not be shown here, but instead will be discussed in the context of the results shown in these two figures. Eigenvectors for shorter wavelength and higher oscillatory mode perturbations, for instance, exhibited a marked preference for an increasing number of vortices located in the interior of the cylinder. The lowest instability mode when $\Lambda = 0$, for instance, exhibited unicellular motion below $k = 2.97$, as was shown in Fig. 6, but bicellular motion with two counter rotating vortices for each wavelength segment of the cylinder when $k \geq 2.97$. As the value of Λ was increased, these transitions occurred at progressively shorter wavelengths, thus the addition of electric current seemed to suppress the emergence of multicellular flows. Distending perturbations long in wavelength exhibited unicellular flow fields similar to the displacement vectors depicted in Fig. 6. Shorter wavelength disturbances, on the other hand, exhibited a multicellular flow field, such as was depicted in Fig. 7. The emergence of these counter-rotating vortices apparently has a stabilizing influence on distending perturbations, since they cause a redistribution of material back into necked regions. It is therefore significant that the value of k where the flow transition occurred was nearly coincident with the solution for the most unstable wavelength depicted in Fig. 5.

Velocity and density profiles at the long wavelength limit for plastic flow were also determined. Solutions for oscillatory perturbations could evidently be separated into two categories: solutions where v_r was zero and no displacement in the radial boundary occurred, and solutions v_z was zero and no change in density occurred. The addition of electric current altered only the solutions where v_r was nonzero; moreover, these solutions were identical to the results attained for purely elastic disturbances. For distending disturbances, a single solution for ω_r existed for each value of Λ, and the associated eigenvectors possessed zero axial velocities.

Since many of the features observed in the plastic flow solutions were also evident in purely elastic solutions, the underlying physical mechanisms contributing to the observed phenomena obviously exhibit distinct similarities. However, unlike the elastic perturbations, distending instabilities were achieved for plastic flow at all levels of current, and the frequencies of oscillatory perturbations were significantly suppressed. Although the condition of plastic flow obviously affects all components of the stress tensor in the cylinder, apparently the distinctions mentioned above can be understood by considering the consequences only on the axial stress. For instance, it is evident from Eq. (20) that the first-order deviatoric axial stress vanishes when plastic flow is initiated in the cylinder. However, as was identified for the elastic disturbances in Part I of this paper, it is the distortional component of σ_{z1} that contributes the restoring tendencies to the cylinder; the mean part of σ_{z1} tends to interfere with the restoring stress gradient since it is compressive in necked regions. Apparently, the absence of this restoring stress gradient results in an unconditional instability, and distending perturbations are achievable at all levels of current. Larger currents enhance this instability and result in even larger growth rates, since the increase in magnetic pressure in necked regions of the cylinder tends to force even more material out of the neck. Moreover, the vanishing deviatoric axial stress gradient apparently accounts for the frequency repression in the oscillatory perturbations, since the axial stress does not contribute to restoring tendencies in the cylinder for the plastic flow condition.

6 Conclusions

The motion of elastic-plastic cylinders is considerably influenced by the presence of electromagnetic fields. When a large axial electric current is introduced and the magnetic pressure exceeds a value of about two thirds of the bulk modulus, a critical condition is achieved and the radius of the cylinder races down very rapidly to a small value. Moreover, electromagnetic fields can exert a substantial influence on both the oscillatory and distending instability spectrums. Azimuthal magnetic fields have a tendency to restrain the frequency of perturbations that fluctuate in amplitude with time, and hasten the growth rate of instabilities that grow in magnitude over time. Both of these tendencies appear to be intensified once the yield criterion for the cylinder has been reached. These results qualitatively agree with the breakup phenomenon observed in thin metallic wires (Chace and Moore, 1968); however, the inclusion of thermoelastic-plastic effects and a rigorous treatment of large strains will be needed before any quantitative comparisons can be made.

Results from this analysis can be applied to show how electromagnetic fields might be used to alter or enhance the instabilities in ductile metal cylinders. For example, if electromagnetic fields are to be used to initiate disruptive instabilities in cylinders, the analysis for purely elastic disturbances in the accompanying paper indicates that the magnetic pressure must be on the order of the bulk modulus of the material. The electric current required to meet this criterion is unrealistically large and on the order of 30–40 MA for a 2 cm diameter cylinder composed of materials such as steel or tungsten. However, thermal softening of the cylinder occurs due to ohmic heating and would significantly reduce the magnitude of the current required, since the bulk modulus tends to decrease with temperature. More importantly, however, is the realization that plastic flow can be initiated when the magnetic pressure is on the order of the yield stress, requiring a lower current of 2–3 MA for a 2 cm cylinder of steel or tungsten. Although this current is still quite large, it can be achieved using state-of-the-art pulsed power supplies. Moreover, since the yield strength of these metals decreases with temperature, and essentially vanishes when the melting temperature is exceeded, ohmic heating may significantly increase the growth rate of instabilities in the cylinder. An investigation of the thermal effects is already underway.

Acknowledgments

The technical work described in this paper was funded by the U. S. Army Research Office under contract No. DAAL03-92-K-0005, whose support is gratefully acknowledged.

References

Bleich, H. H., and Nelson, I., 1966, "Plane Waves in Elastic-Plastic Half Space Due to Combined Surface Pressure and Shear," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 149–158.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Chace, W. G., and Moore, H. K., ed., 1968, *Exploding Wires,* Plenum, New York.

Chou, P. C., and Carleone, J., 1977, "The Stability of Shaped-charge Jets," *Journal of Applied Physics,* Vol. 48, No. 10, pp. 4187–4195.

Graff, K. F., 1975, *Wave Motion in Elastic Solids,* Ohio State University Press, Columbus, OH.

Kruskal, M., and Schwartzschild, M., 1954, "Some Instabilities of a Completely Ionized Plasma," *Proceedings of the Royal Society of London,* Vol. A223, pp. 348–360.

Littlefield, D. L., and Powell, J. D., 1990, "The Effect of Electromagnetic Fields on the Stability of a Uniformly Elongating Plastic Jet," *Physics of Fluids A,* Vol. 2, No. 12, pp. 2240–2248.

Littlefield, D. L., 1991, "Finite Conductivity Effects on the MHD Instabilities in Uniformly Elongating Plastic Jets," *Physics of Fluids A,* Vol. 3, No. 6, pp. 1666–1672.

Littlefield, D. L., 1994, "Thermomechanical and Magnetohydrodynamic Stability of Elongating Plastic Jets," *Physics of Fluids,* Vol. 6, No. 8, pp. 2722–2729.

Littlefield, D. L., 1995, research in progress.

Littlefield, D. L., 1996a, "Magnetomechanical Instabilities in Elastic-Plastic Cylinders, Part I: Elastic Response," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 734–741.

Littlefield, D. L., 1996b, "Complex Solutions to Higher Order Equations: Magnetomechanical Stability of Elastic-Plastic Cylinders," in preparation.

Maugin, G. A., 1988, *Continuum Mechanics of Electromagnetic Solids,* Elsevier, New York.

Moon, F. C., 1984, *Magneto-solid Mechanics,* John Wiley and Sons, New York.

Nowacki, W. K., 1978, *Stress Waves in Non-elastic Solids,* Pergamon Press, Oxford, UK.

Pack, D. C., 1988, "On the Perturbation and Breakup of a High-speed, Elongating Metal Jet," *Journal of Applied Physics,* Vol. 63, No. 6, pp. 1864–1871.

Rayleigh, 1877, "The Theory of Sound," *London Mathematical Society Proceedings,* Vol. 17 (Dover Reprint, New York, 1945).

Romero, L. A., 1989, "The Instability of Rapidly Stretching Plastic Jets," *Journal of Applied Physics,* Vol. 65, No. 8, pp. 3006–3016.

Yarin, A. L., 1994, "On Instability of Rapidly Stretching Metal Jets Produced by Shaped Charges," *International Journal of Engineering Science,* Vol. 32, No. 5, pp. 847–862.

APPENDIX

Dispersion Relations for the Plastic Response

Solutions to the plastic response in the cylinder are given in Eqs. (23)–(24). The arbitrary constants c_1 and c_2 appearing in these equations are determined from substitution of the boundary conditions, given in Eqs. (25)–(26). This substitution yields two linear homogeneous algebraic equations, having nontrivial solutions only when their determinant is zero. The form of this determinant depends on the roots of a characteristic polynomial for β, given in Eq. (22). When β^2 is real, the determinant is given by

$$\left[-\frac{k^2Z(1-\Psi)}{\Lambda}+\frac{\omega^2+\Phi k^2}{\Omega}+\beta_1^2\right]\times\left(\Phi-\frac{\omega^2+\Phi k^2}{\Omega\beta_2^2}\right)\beta_2 J_1(\beta_1)J_0(\beta_2)$$

$$-\left[-\frac{k^2Z(1-\Psi)}{\Lambda}+\frac{\omega^2+\Phi k^2}{\Omega}+\beta_2^2\right]\times\left(\Phi-\frac{\omega^2+\Phi k^2}{\Omega\beta_1^2}\right)\beta_1 J_0(\beta_1)J_1(\beta_2)$$

$$+(2\Phi+Z)(\beta_2^2-\beta_1^2)J_1(\beta_1)J_1(\beta_2)=0, \quad \text{(A1)}$$

where β_1^2 and β_2^2 are the two roots for β^2, and $Z = 4\Lambda/(4 + \Lambda - 2\Phi\Psi/3)$. When β^2 is imaginary or complex, the determinant is given as

$$\left\{\left[\frac{k^2Z(1-\Psi)}{\Lambda}-\frac{\omega^2+\Phi k^2}{\Omega}+\beta_m^2\right]\Phi+\left[\frac{k^2Z(1-\Psi)}{\Lambda}-\frac{\omega^2+\Phi k^2}{\Omega}-(3\beta_r^2-\beta_i^2)\right]\frac{\omega^2+\Phi k^2}{\beta_m^2\Omega}\right\}\times\beta_i[RJ_0(\beta)RJ_1(\beta)+IJ_0(\beta)IJ_1(\beta)]$$

$$+\left\{\left[\frac{k^2Z(1-\Psi)}{\Lambda}-\frac{\omega^2+\Phi k^2}{\Omega}-\beta_m^2\right]\Phi-\left[\frac{k^2Z(1-\Psi)}{\Lambda}-\frac{\omega^2+\Phi k^2}{\Omega}+(3\beta_r^2-\beta_i^2)\right]\frac{\omega^2+\Phi k^2}{\beta_m^2\Omega}\right\}\times\beta_r[RJ_1(\beta)IJ_0(\beta)-IJ_1(\beta)RJ_0(\beta)]$$

$$-2\beta_r\beta_i(2\Phi+Z)\{[RJ_1(\beta)]^2+[IJ_1(\beta)]^2\}=0, \quad \text{(A2)}$$

where β_r and β_i are moduli of the real and imaginary components of β, respectively; and $\beta_m^2 = \beta_r^2 + \beta_i^2$. Solutions to these dispersion relations identify the (ω, k) regions where solutions to the eigenvalue problem exist.

When the wavelength is very long, the dispersion relation can be simplified. The expression is

$$I_1\left(\frac{\omega}{\sqrt{\Phi}}\right)\left[\omega\sqrt{\Omega}I_0\left(\frac{\omega}{\sqrt{\Omega}}\right)-(2\Phi+Z)I_1\left(\frac{\omega}{\sqrt{\Omega}}\right)\right]=0. \quad \text{(A3)}$$

Similarly, the dispersion relation for the incipient instability condition becomes

$$\left[\frac{Z(1-\Psi)}{\Lambda}\left(1+\frac{1}{\sqrt{\Omega}}\right)-1\right]\times\Phi\beta_i[RJ_0(\beta)RJ_1(\beta)+IJ_0(\beta)IJ_1(\beta)]+\left[\frac{Z(1-\Psi)}{\Lambda}\left(1-\frac{1}{\sqrt{\Omega}}\right)-1\right]\times\Phi\beta_r[RJ_1(\beta)IJ_0(\beta)-IJ_1(\beta)RJ_0(\beta)]$$

$$-\frac{\sqrt{3\Omega+\Phi}}{2\Omega}(2\Phi+Z)\{[RJ_1(\beta)]^2+[IJ_1(\beta)]^2\}=0, \quad \text{(A4)}$$

where $\beta_r^2 = (1 + 2\sqrt{\Omega})k^2/4\Omega$ and $\beta_i^2 = (2\sqrt{\Omega} - 1)k^2/4\Omega$.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Scaled Memory Description of Hysteretic Material Behavior

J. P. Bardet
Associate Professor,
Civil Engineering Department,
University of Southern California,
Los Angeles, CA 90089-2531

Herein the concept of scaled memory (SM) is proposed to describe hysteretic material behaviors. SM scales the nonlinear variation of tangential modulus during monotonic loadings into a piecewise linear distribution, which can easily be modified to simulate specific characteristics of material behavior during cyclic loadings. As an example, SM is applied to eliminate artificial ratchetting in bounding surface plasticity, and to produce closed stress-strain loops during cyclic loadings, without introducing a single material constant. SM is generalized to six dimensions in the particular case of pressure-independent materials, and is applied to simulate the hysteretic responses of metals.

Introduction

Since 1967, several constitutive frameworks have been proposed to model the cyclic behavior of materials (Chaboche, 1986)—multiple-yield surface plasticity (Iwan, 1967; Mróz, 1967), bounding surface plasticity (Dafalias and Popov 1975; Krieg, 1975), endochronic theory (Valanis, 1971), subloading-surface elastoplasticity (Hashiguchi, 1980), and discrete memory (Mróz and Norris, 1982; Wack et al., 1983). Chu (1984, 1987) used a continuous field of yield surface similar to Mróz and Norris (1982), Mróz and Lind (1975), and Mróz (1984), and predicted well the multidimensional cyclic stress-strain relationships of metals which are experimentally measured. However, Chu conceded the need for an excessive memory capacity in numerical analyses for cyclic loadings of diminishing amplitude. Hashiguchi (1993) discussed the modeling of cyclic ratchetting, a phenomenon which is characterized by the cyclic accumulation of deformation. He pointed out that bounding surface plasticity (e.g., Dafalias and Herrmann, 1982; Bardet, 1986) produces excessively large ratchetting. Attempts to mitigate this artificial ratchetting by the introduction of sophisticated mapping rules proved not practical (e.g., Wang et al., 1990; Abu-Lebdeh and Voyiadjis, 1993). These complications destroy the very mathematical simplicity which gives bounding surface plasticity its appeal. Constitutive frameworks other than bounding surface plasticity (e.g., multiple yield surface plasticity) may not display artificial ratchetting. However, the large number of yield surfaces they require results in numerous material constants, and makes them more cumbersome to use for practicing engineers.

There is a definite need for a practical and realistic description of hysteretic behavior in the fields of geomechanics and soil dynamics. During earthquakes, soils and rocks are subjected to cyclic loadings which are substantially different from those encountered in other engineering fields. These cyclic loadings are multidimensional, have a significant static offset caused by gravity, and contain all amplitude ranges activating various features of material behavior from elastic response to complete failure.

In an attempt to improve the description of cyclic behaviors, we propose the general concept of scaled memory (SM). As an illustration, the SM will be applied to eliminate the artificial ratchetting of bounding surface plasticity. The SM concept will first be defined in one dimension, then generalized in six dimensions for the particular case of pressure-independent materials.

Scaled Memory During One-Dimensional Monotonic Loadings

Introduction of Basic SM Concept. The basic SM concept is introduced by considering the hyperbolic model of Kondner (1963) which provides us with an example of stress-strain curves during monotonic (i.e., without a loading-reversal) uniaxial compression:

$$\sigma = \frac{\epsilon}{\frac{1}{S_0} + \frac{\epsilon}{\sigma_{\max}}} \tag{1}$$

where S_0 is the initial slope at $\sigma = 0$, and $\sigma_{\max}$ is the compressive failure stress. The stress-strain curve during uniaxial tension is obtained from Eq. (1) by replacing $\sigma_{\max}$ with the tensile failure stress $\sigma_{\min}$. Hereafter, stress and strain are positive in compression, and negative in tension. As shown in Fig. 1(a), the initial state is strain and stress-free (i.e., $\epsilon = 0$ and $\sigma = \sigma_0 = 0$). Curves OD and OE represent the stress-strain response curves during monotonic compression and tension, respectively. Eq. (1) may be written incrementally:

$$d\sigma = S d\epsilon \tag{2}$$

where $d\sigma$ is the stress increment, $d\epsilon$ is the strain increment, and S is the variable slope of the stress-strain curve. By differentiating Eq. (1), S is not only a function of stress σ,

$$S(\sigma) = S_0\left(1 - \frac{\sigma}{\sigma_{\max}}\right)^2, \tag{3}$$

but also a function of the dimensionless distance δ between σ and the failure stresses $\sigma_{\min}$ and $\sigma_{\max}$,

$$S(\sigma) = S(\delta) = S_0\delta^2, \tag{4}$$

where δ is

$$\delta = \begin{cases} 1 - \dfrac{\sigma}{\sigma_{\max}} & \text{if} \quad \sigma \geq 0 \\ 1 - \dfrac{\sigma}{\sigma_{\min}} & \text{if} \quad \sigma < 0. \end{cases} \tag{5}$$

As shown in Fig. 1(b), S is a nonlinear function of σ which also depends on material properties (e.g., S_0), whereas, as shown in

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, July 14, 1994; final revision, Aug. 1, 1995. Associate Technical Editor: J. W. Drugan.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

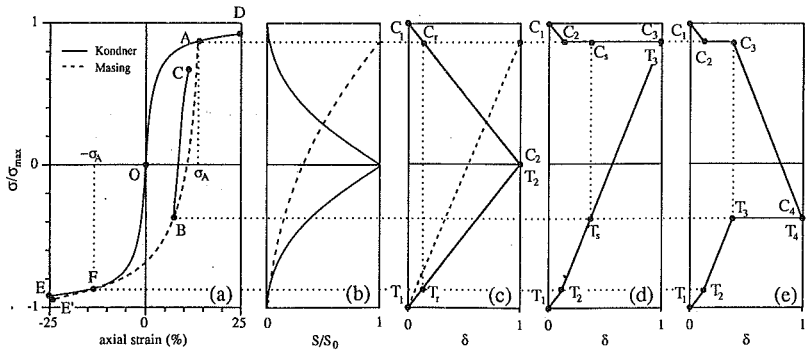


Fig. 1 (*a*) Stress-strain response curves simulated by Kondner during compression and tension, and by Masing and SM during unloading from A, (*b*) corresponding variations of tangential modulus, (*c*) scaled memory at point O, (*d*) scaled memory after unloading at point A, and (*e*) scaled memory after reloading at point B

Fig. 1(c), δ is a bilinear function of σ that depends only on the initial stress σ_0 and the failure stresses $\sigma_{\max}$ and $\sigma_{\min}$.

Definitions. Hereafter, δ is called the *scaled modulus.* δ is normalized so that $\delta = 0$ when $\sigma = \sigma_{\max}$ and $\sigma = \sigma_{\min}$; and $\delta = 1$ when $\sigma = \sigma_0 = 0$. The function $S(\delta)$ that scales the tangential modulus S into the scaled modulus δ is called the *scaling function.* For instance, Eq. (4) is the scaling function which scales S from the σ-segment $[0, S_0]$ onto the dimensionless δ-segment [0,1]. The distribution of the scaled modulus, which is represented by points C_1, C_2, T_1, and T_2 in Fig. 1(c), is called the *scaled memory* (SM). By using an index notation, points C_1, C_2, T_1, and T_2 have for coordinates (c_1, δ_1), (c_2, δ_2), (t_1, δ_1), and (t_2, δ_2), respectively, where

$$c_1 = \sigma_{\max}, \quad t_1 = \sigma_{\min}, \quad c_2 = t_2 = \sigma_0,$$
$$\delta_1 = 0, \quad \text{and} \quad \delta_2 = 1. \tag{6}$$

The c_i's and C_i's are for compression while the t_i's and T_i's are for tension. SM is conveniently described by defining the *dipoles* D_1 and D_2, each dipole D_i combining points C_i and T_i, and having coordinates (c_i, t_i, δ_i). The points C_i and T_i generate the *compressive and tensile memory branch,* respectively. As shown in Fig. 1(c), SM initially has two dipoles $D_1(c_1, t_1, \delta_1)$ and $D_2(c_2, t_2, \delta_2)$, and its compressive and tensile memory branches are C_1C_2 and T_1T_2, respectively.

In order to describe more complicated distributions of the scaled modulus, we generalize SM from a bilinear distribution with two dipoles $D_i(c_i, t_i, \delta_i)$ to a piecewise linear distribution with m dipoles. The number m characterizes the memory size. The compressive and tensile memory branches are made of m points $C_i(c_i, \delta_i)$ and $T_i(t_i, \delta_i)$.

SM Description of Observed Behavior. The SM concept was introduced for the particular case $\sigma_{\min} = -\sigma_{\max}$, but it also applies to $\sigma_{\min} \neq -\sigma_{\max}$. In this case, SM becomes unsymmetric as shown in Fig. 2(b). SM may also have more than two

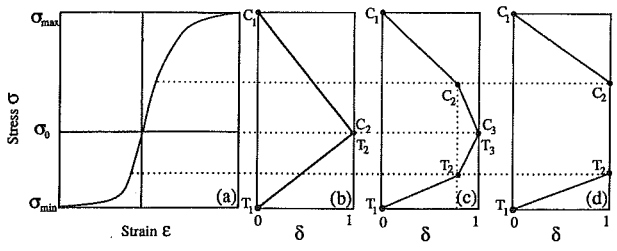


Fig. 2 (*a*) Unsymmetric stress-strain curve during compression and tension, (*b*) scaled memory with two dipoles, and (*c*) scaled memory with three dipoles for bounding surface plasticity without an elastic nucleus. (*d*) Scaled memory with two dipoles for two-surface bounding surface plasticity.

dipoles. For instance, it has three dipoles in Fig. 2(c). The third dipole allows us to describe stress-strain curves that have different shapes in compression and tension with the same scaling function $S(\delta)$. We do not rule out the addition of other dipoles for extremely nonlinear stress-strain curves, but based on our experience (Bardet, 1995), three dipoles are sufficient to fit the stress-strain response curves of most materials.

SM Description of Bounding Surface Plasticity Without Elastic Nucleus. Bounding surface (BS) plasticity (Dafalias and Popov, 1975; Krieg, 1975), like classical elastoplasticity (Hill, 1950), assumes that the incremental strain $d\epsilon$ is the sum of an elastic strain $d\epsilon^e$ and a plastic strain $d\epsilon^p$:

$$d\epsilon = d\epsilon^e + d\epsilon^p = \frac{d\sigma}{E} + \frac{d\sigma}{H} \tag{7}$$

where E is the elastic Young's modulus, and H is the plastic modulus. For instance, H may be selected as a function of δ:

$$H = H(\delta) = h_0 \frac{E\delta^r}{(1-\delta)^s} \tag{8}$$

where h_0, r, and s are material constants, and δ is given by Eq. (5). By using Eq. (7), the tangential modulus S becomes a function of δ:

$$\frac{1}{S(\delta)} = \frac{1}{E} + \frac{1}{H(\delta)}. \tag{9}$$

Equation (8) is a particular definition of $H(\delta)$, which does not preclude the consideration of other expressions. Equation (8) describes the gradual transition from elastic to perfectly plastic responses. When $\sigma = 0$, $H \to \infty$ and the strain predicted by Eq. (7) is purely elastic. Therefore, the elastic deformation is confined to a particular stress, and there is no elastic nucleus. When $\sigma = \sigma_{\max}$ or $\sigma = \sigma_{\min}$, $H = 0$ and the strain is infinitely large, which corresponds to failure. Eqs. (8) and (9) provide us with the scaling function that scales S from the σ-segment $[0, E]$ onto the dimensionless δ-segment [0, 1]. The material constants h_0, r and s allow us to control the growth of plastic behavior. Their effects are discussed by Bardet (1990), and Dafalias and Herrmann (1982).

When the elastoplastic description of Eq. (7) is used for Kondner's model, Young's modulus E is equal to S_0, and the plastic modulus H is

$$H = H(\delta) = E \frac{\delta^2}{1-\delta^2}. \tag{10}$$

Equation (10) produces a similar variation for H as Eq. (8), although it is not strictly encompassed by Eq. (8). The BS

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

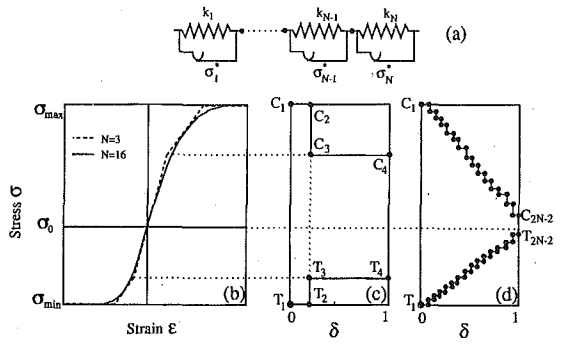


Fig. 3 Multiple yield surface plasticity: (*a*) one-dimensional representation with series of spring-slider elements; (*b*) simulated stress-strain response curve, (*c*) scaled memory for three elements, and (*d*) scaled memory for 16 elements

model of Eq. (8) and Kondner's models have different scaling functions but identical scaled memories.

SM Description of Two-Surface Bounding Surface Plasticity. The two-surface version of bounding surface plasticity (Krieg, 1975) has a yield surface which translates and/or expands within the bounding surface. The response is elastic inside the yield surface. The previous BS model is a particular of two-surface model with a vanishing elastic range. As shown in Fig. 2(*d*), the SM of the two-surface BS has three dipoles. On segment C_2T_2, the scaled modulus δ is equal to 1 and the tangential modulus S is equal to Young's modulus E. Therefore, the deformation is purely elastic between C_2 and T_2, which describes the elastic nucleus of a two-surface BS. The scaling relations used for the SM of Figs. 2(*c*) and 2(*d*) may be of the same type (e.g., Eqs. (8) and (9)), but must obviously have different material constants (e.g., h_0, r, and s) to produce similar stress-strain curves.

SM Description of Multiple Yield Surface Plasticity. Multiple yield surface plasticity (MYSP) with linear hardening was proposed to simulate hysteretic behavior (Iwan, 1967; and Mróz, 1967). As shown in Fig. 3(*a*), in the one-dimensional case, MYSP is represented by a series of N spring-slider elements. The spring value of the ith element is k_i, and its compressive yield stress is σ_i^*. The spring values form a decreasing series, and the yield stresses form an increasing series (i.e., $k_1 > k_2 > \ldots > k_N$, and $\sigma_1^* < \sigma_2^* < \ldots < \sigma_N^*$). σ_1^* and k_N are set equal to zero to obtain an elastic response initially, and a perfectly plastic response at large strain. Such a model simulates a piecewise linear stress-strain curve with an abrupt change in slope S when a slider yields. When j sliders have yielded (i.e., when $\sigma_j^* \leq \sigma < \sigma_{j+1}^*$), then the slope S is equal to S_j, where

$$\frac{1}{S_j} = \sum_{i=1}^{j} \frac{1}{k_i}. \tag{11}$$

The slope S takes N discrete values S_j decreasing from k_1 to 0 during compression and tension. These values S_j can be scaled by using a scaling function similar to that used for bounding surface plasticity, which produces N scaled moduli δ_j^*. As shown in Figs. 3(*c*) and 3(*d*), the scaled memory of MYSP is a staircase. The coordinates of dipole D_i (c_i, t_i, δ_i), the scaled modulus δ_j^*, and the yield stresses σ_i^* of the ith slider are related through

$$c_1 = c_2 = -t_1 = -t_2 = \sigma_N^*,$$

$$c_3 = c_4 = -t_3 = -t_4 = \sigma_{N-1}^*, \ldots, \quad \text{and}$$

$$c_{2N-3} = c_{2N-2} = -t_{2N-3} = -t_{2N-2} = \sigma_2^*$$

$$\delta_1 = \delta_N^* = 0, \ \delta_2 = \delta_3 = \delta_{N-1}^*, \quad \delta_4 = \delta_5 = \delta_{N-2}^*, \ldots,$$

$$\text{and} \quad \delta_{2N-2} = \delta_1^* = 1. \tag{12}$$

Therefore, the total number of dipoles is equal to $2(N\text{-}1)$, and the SM staircase has N-1 steps. As shown in Fig. 3(*c*), a three-element MYSP model (i.e., $N = 3$) has four dipoles but three different δ values, and simulates a piecewise linear stress-strain curve with three different slopes. As shown in Figs. 3(*b*) and 3(*d*), the higher the number N of sliders (yield surfaces), the smoother the piecewise linear stress-strain curve of MYSP becomes, and the more accurate the simulation of the actual measured response could become. When $N \to \infty$, the SM staircase of MYSP converges toward a smooth δ-distribution described by an infinite number of dipoles. This continuous δ-distribution, although very accurate, would be impractical. In the SM framework, however, it can be approximated by using a piecewise linear distribution with a finite number of dipoles. Therefore SM provides us with a more realistic and practical description of stress-strain curves than MYSP. Additional comparison between SM and MYSP models can be found in Bardet (1995).

Remarks. SM provides a geometrical representation of the stress-strain responses which are measured in the laboratory or simulated by elastoplastic models. SM is controlled by the scaling function between the tangential and scaled moduli. Different scaling functions may produce similar stress-strain curves, but different SM. It is recommended to carefully select the scaling function in order to have a SM that is initially simple (e.g., bilinear). The lack of uniqueness for the scaling function does not impede the SM capabilities, but illustrates its versatility to describe stress-strain response curves.

Hereafter, we assume that S is only a function of the scaled modulus δ, and exclude the effects of other internal variables. We also rule out strain softening as the result of internal variables. The introduction of internal variables and strain softening is beyond the scope of this paper, and will be investigated in the future.

SM was introduced by using a stress-strain curve with a slope $S(\sigma)$ that vanishes when σ is equal to $\sigma_{\min}$ or $\sigma_{\max}$. However, it also applies to strain-hardening stress-strain curves for which $S(\sigma)$ never vanishes, even at large strain. In this case, Eq. (3) can be easily modified so that $S(\sigma)$ becomes practically constant to some small positive value when σ exceeds $\sigma_{\min}$ or $\sigma_{\max}$.

Scaled Memory During One-Dimensional Cyclic Loadings

The SM concept becomes most useful when modeling cyclic responses. Its piecewise-linear structure can be modified by many basic geometrical transformations, free of the difficulties associated with nonlinearities and material parameters. The SM benefits are illustrated by requiring that the simulated cyclic curves (1) obey Masing's similitude, (2) do not overshoot monotonic responses, and (3) remain closed during cycles of constant amplitude. Then we will devise an approximation scheme that controls the SM size. The three above requirements reflect experimental observations obtained for most materials, but do not preclude us from considering other requirements on material responses or from generating other evolutions for SM in the future.

SM During Unloading. For an unloading (i.e., $d\sigma < 0$) about stress σ_A and strain ϵ_A, Masing (1929) postulated that the stress-strain response is

$$\sigma_A - \sigma = 2f\left(\frac{\epsilon_A - \epsilon}{2}\right) \tag{13}$$

where the function f defines the stress-strain curve during loading (e.g., Eq. (1)). Masing assumed that the compressive and

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

tensile stress-strain curves are symmetric (e.g., $\sigma_{\min} = -\sigma_{\max}$ in Eq. (1)). As shown in Fig. 1(a) the unloading curve ABFE′ given by Eqs. (1) and (13) intersects the tensile curve OFE at point F (i.e., $\sigma = -\sigma_A$), and overshoots it after point F (i.e., $\sigma < -\sigma_A$). Equation (13) implies that the unloading slope $S'(\sigma)$ and loading slope $S(\sigma)$ are related through

$$S'(\sigma) = S\left(\frac{\sigma_A - \sigma}{2}\right). \tag{14}$$

After unloading at A, $S'(\sigma)$ is shown in Fig. 1(b). By using Eqs. (5) and (14), the corresponding scaled modulus δ, as shown in Fig. 1(d), becomes

$$\delta = 1 + \frac{\sigma - \sigma_A}{2\sigma_{\max}}. \tag{15}$$

We prevent the curve ABFE′ generated by Masing's similitude from overshooting the tensile curve OFE by setting it equal to the tensile curve FE after point F. Point F is determined by assuming identical slopes at points F and A, which will guarantee the closure of stress-strain loops during constant amplitude cycles as shown later. As shown in Fig. 1(c), point A corresponds to point C_r (c_r, δ_r) on the compressive memory branch (i.e., $c_r = \sigma_A$, and $\delta_r = \delta_A$). Point F corresponds to the point $T_r(t_r, \delta_r)$ on the tensile memory branch, which has the same scaled and tangential moduli as C_r. In this particular case, $t_r = \sigma_F = -\sigma_A$ because the compressive and tensile memory branches are symmetric. The corrected unloading curve ABFE is easily obtained by modifying Eq. (15):

$$\delta = \begin{cases} 1 + \dfrac{\sigma - \sigma_A}{2\sigma_{\max}} & \text{when} \quad -\sigma_A \le \sigma \le \sigma_A \\ 1 - \dfrac{\sigma}{\sigma_{\min}} & \text{when} \quad \sigma_{\min} \le \sigma \le -\sigma_A. \end{cases} \tag{16}$$

As shown in Fig. 1(d), SM has three dipoles (i.e., $m = 3$) during unloading. The new dipole $D_r(c_r, t_r, \delta_r)$ is inserted into the SM of Fig. 1(c) so that the δ_i's generate a strictly decreasing series (i.e., $\delta_3 > \delta_2 > \delta_1$). Whereas dipole D_1 is unchanged, D_2 is set equal to D_r, and D_3 becomes

$$(c_3, t_3, \delta_3) = (c_r, c_r, 1). \tag{17}$$

By using an index notation, the unloading response ABFE is calculated incrementally by using

$$\delta = \begin{cases} \delta_2 + \dfrac{\delta_3 - \delta_2}{t_3 - t_2}(\sigma - t_2) & \text{when} \quad t_3 \le \sigma \le t_2 \\ \delta_1 + \dfrac{\delta_2 - \delta_1}{t_2 - t_1}(\sigma - t_1) & \text{when} \quad t_2 \le \sigma \le t_1. \end{cases} \tag{18}$$

Equation (18) not only generates part of the unloading curve with Masing's similitude, but it also generalizes Masing's similitude to unsymmetric response curves with different compressive and tensile failure stresses (i.e., $\sigma_{\min} \neq -\sigma_{\max}$). Again, SM guarantees Masing's similitude and no tensile overshooting, regardless of the scaling function and material properties.

SM During Arbitrary Cyclic Loadings. The SM definition will now be extended to reloading (i.e., $d\sigma > 0$). As shown in Fig. 1(d), when reloading takes place at B, the dipole D_s (c_s, t_s, δ_s) at B has for coordinates $t_s = \sigma_B$, $\delta_s = \delta_B$ and

$$c_s = \begin{cases} c_1 + \dfrac{c_2 - c_1}{\delta_2 - \delta_1}(\delta_s - \delta_1) & \text{when} \quad t_1 \le \sigma_B \le t_2 \\ c_2 + \dfrac{c_3 - c_2}{\delta_3 - \delta_2}(\delta_s - \delta_2) & \text{when} \quad t_3 \le \sigma_B \le t_2. \end{cases} \tag{19}$$

Table 1 Stress-strain calculation algorithm for SM models

initialize stress and strain: $\varepsilon=0$, $\sigma=\sigma_0$, $\delta=1$

initialize memory: $m=2$, $c_1=\sigma_{max}$, $t_1=\sigma_{min}$, $\delta_1=0$, $c_2=\sigma_0$, $t_2=\sigma_0$, $\delta_2=1$

(*) *update memory (only if loading is reversed, i.e., $d\sigma$ changes sign)*

if $d\sigma>0$ then find $i=1,..,m-1$ $\ni$ $t_i<\sigma\le t_{i+1}$,

$$m \leftarrow i+2,\ t_{m-1} \leftarrow \sigma,\ c_{m-1} \leftarrow c_i + \frac{c_{i+1}-c_i}{\delta_{i+1}-\delta_i}(\delta - \delta_i),\ \delta_{m-1} \leftarrow \delta$$

elseif $d\sigma<0$ then find $i=1,..,m-1$ $\ni$ $c_{i+1}\le\sigma<c_i$,

$$m \leftarrow i+2,\ c_{m-1} \leftarrow \sigma,\ t_{m-1} \leftarrow t_i + \frac{t_{i+1}-t_i}{\delta_{i+1}-\delta_i}(\delta - \delta_i),\ \delta_{m-1} \leftarrow \delta$$

endif

$c_m \leftarrow \sigma$, $t_m \leftarrow \sigma$, $\delta_m \leftarrow 1$

reduce memory size (only if $m = M+1$)

find $i=2,...,M$ $\ni$ $A_i = minimum(A_2,...,A_M)$ where

$$A_k^2 = \left(t_k - t_{k-1} - \frac{t_{k+1}-t_{k-1}}{\delta_{k+1}-\delta_{k-1}}(\delta_k - \delta_{k-1})\right)^2 + \left(c_k - c_{k-1} - \frac{c_{k+1}-c_{k-1}}{\delta_{k+1}-\delta_{k-1}}(\delta_k - \delta_{k-1})\right)^2$$

$t_j \leftarrow t_{j+1}$, $c_j \leftarrow c_{j+1}$, $\delta_j \leftarrow \delta_{j+1}$ for $j=i,..., M$

$m=M$

calculate stress-strain response

if $d\sigma>0$, then find $i=1,..,m-1$ $\ni$ $c_{i+1}\le\sigma<c_i$, then $\delta = \delta_i + \frac{\delta_{i+1}-\delta_i}{c_{i+1}-c_i}(\sigma - c_i)$

if $d\sigma<0$, then find $i=1,..,m-1$ $\ni$ $t_i<\sigma\le t_{i+1}$, then $\delta = \delta_i + \frac{\delta_{i+1}-\delta_i}{t_{i+1}-t_i}(\sigma - t_i)$

$$d\varepsilon = \frac{d\sigma}{S(\delta)},\ \varepsilon \leftarrow \varepsilon + d\varepsilon,\ \sigma \leftarrow \sigma + d\sigma$$

go to * until loading completed

When $t_3 \le \sigma_B \le t_2$, SM has four dipoles (i.e., $m = 4$). Dipoles D_1 and D_2 are unchanged, D_3 is set equal to D_s, and D_4 becomes

$$(c_4, t_4, \delta_4) = (t_s, t_s, 1). \tag{20}$$

If $t_2 \le \sigma_B \le t_1$, SM would have only three dipoles (i.e., $m = 3$). Dipole D_1 would be unchanged, and D_2 would be set equal to D_s, and D_3 would become

$$(c_3, t_3, \delta_3) = (t_s, t_s, 1). \tag{21}$$

Therefore, the memory size m increases only when the loading is reversed at a stress between t_{m-1} and c_{m-1}. It decreases during large loading reversals because dipole D_k is erased when the loading is reversed at a stress σ outside C_kT_k (i.e., $c_k < \sigma$ or $\sigma < t_k$).

Table 1 generalizes the SM definition from the loading-unloading-reloading cycle OABC of Fig. 1 to arbitrary cycles. The algorithm of Table 1 initializes and updates SM, and calculates the stress-strain response for stress-controlled loadings. It can be easily adapted for strain-controlled cycles. Table 1 confers to SM the following properties:

$$\delta_m = 1 > \delta_{m-1} > \ldots > \delta_{i+1} > \delta_i > \delta_{i-1} > \ldots > \delta_2 > \delta_1 = 0 \tag{22}$$

$$c_1 \ge c_2 \ge \ldots \ge c_{m-1} \ge c_m = t_m \ge t_{m-1} \ge \ldots \ge t_2 \ge t_1. \tag{23}$$

The strict inequalities of Eq. (22) ensure that dipoles are uniquely defined. As stated in Eq. (23), several stresses c_i or t_i may coincide, which corresponds to abrupt variations of the tangential modulus. Table 1 does not apply to MYSP, because MYSP has a fixed number of dipoles, each having a constant value of δ_i. At the loading reversals, MYSP translates the activated dipoles without creating new ones.

As shown in Table 1, the slope S of the stress-strain curve is a function of dipoles. These dipoles are internal variables that vary in a discontinuous way in order to represent the effects of the loading reversal history on the material response.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Closure of Stress-Strain Loops. The closure of stress-strain loops during stress cycles of fixed amplitude is established as follows. After the second cycle between stresses σ_A and σ_B, SM alternates between two symmetric triangles for $\sigma_A \leq \sigma \leq \sigma_B$. The memory size m remains constant, and only the dipoles D_m and D_{m-1} are active. For loading from σ_A to σ_B, the strain increment $\Delta\epsilon_{AB}$ is

$$\Delta\epsilon_{AB} = \int_{\sigma_A}^{\sigma_B} \frac{d\sigma}{S(\sigma)} = \frac{\sigma_B - \sigma_A}{\delta_{m-1} - 1} \int_1^{\delta_{m-1}} \frac{d\delta}{S(\delta)}, \quad (24)$$

because σ and δ are related through

$$\sigma = \sigma_A + \frac{\sigma_B - \sigma_A}{1 - \delta_{m-1}} (1 - \delta). \quad (25)$$

For reloading from σ_B to σ_A, the strain $\Delta\epsilon_{BA}$ is

$$\Delta\epsilon_{BA} = \int_{\sigma_B}^{\sigma_A} \frac{d\sigma}{S(\sigma)} = \frac{\sigma_A - \sigma_B}{\delta_{m-1} - 1} \int_1^{\delta_{m-1}} \frac{d\delta}{S(\delta)} = -\Delta\epsilon_{AB}, \quad (26)$$

because

$$\sigma = \sigma_B + \frac{\sigma_A - \delta_B}{1 - \delta_{m-1}} (1 - \delta). \quad (27)$$

Therefore, Eq. (26) implies that the stress-strain loop is closed during stress cycles, regardless of the material properties, and rules out the artificial ratchetting observed in bounding surface plasticity.

Approximation of Scaled Memory. As previously stated, the number of dipoles remains small when there are large loading reversals. However, the number of dipoles may excessively increase for stress cycles that have a slowly decreasing amplitude. Such stress cycles are not hypothetical, but are commonly encountered in earthquake engineering, when seismic accelerations slowly decay with time. Again, we will take advantage of the piecewise linear SM structure to control the number of dipoles.

We propose to control the memory size by approximating SM with M dipoles maximum. When $m > M$, we suggest the elimination of the dipole D_k that least changes SM, but exclude from the elimination process the first and last dipoles, which correspond to the last loading reversal and the failure stresses. We create the interpolated dipole $D_p(c_p, t_p, \delta_k)$ from dipoles D_{k-1} and D_{k+1} and from the scaled modulus δ_k of dipole D_k:

$$t_p = t_{k-1} - \frac{t_{k+1} - t_{k-1}}{\delta_{k+1} - \delta_{k-1}} (\delta_k - \delta_{k-1}), \quad \text{and}$$

$$c_p = c_{k-1} - \frac{c_{k+1} - c_{k-1}}{\delta_{k+1} - \delta_{k-1}} (\delta_k - \delta_{k-1}). \quad (28)$$

Because dipole D_p is linearly interpolated from dipoles D_{k-1} and D_{k+1}, it does not influence the piecewise linear memory branches for the scaled memory that has no dipole D_k. Therefore, the dipole D_k that least contributes to SM is the closest to dipole D_p, and corresponds to the minimum value of A_k where

$$A_k^2 = (t_k - t_p)^2 + (c_k - c_p)^2. \quad (29)$$

The memory contraction algorithm is shown in Table 1. The SM performance will later be compared for particular examples.

Like Mróz (1984) and Chu (1984) approaches, the SM concept uses a discrete memory and modifies it at stress reversals. However, unlike Mróz and Chu approaches, the SM concept scales the nonlinear variation of plastic modulus into a piecewise linear distribution of scaled modulus. Such a scaling is a major departure from other models. For instance, it allows us to control the number of memory surfaces, even for cyclic loadings of diminishing amplitude. The scaled memory framework encompasses not only the discrete memory of Mróz and Chu but also bounding surface plasticity.

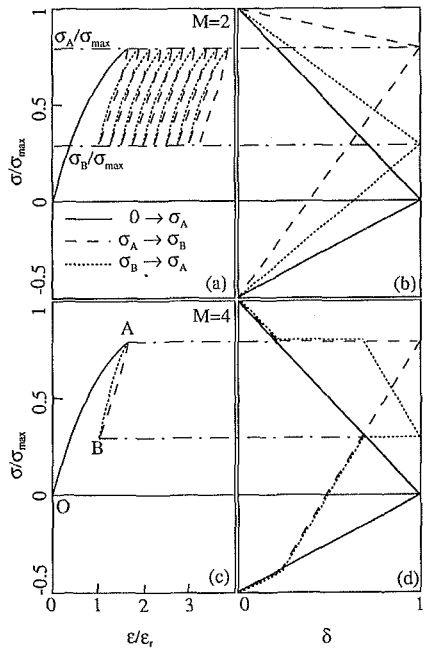


Fig. 4 Stress-strain response curve simulated by (*a*) original ($M = 2$) and (c) improved ($M \geq 4$) bounding surface plasticity model, and corresponding scaled memories (*b* and *d*) during unsymmetric stress cycles

Application to Bounding Surface Plasticity. We apply the SM concept to eliminate the artificial ratchetting of bounding surface plasticity mentioned by Hashiguchi (1993), but do not claim that cyclic ratchetting, which is experimentally observed on metals (e.g., Hassan and Kyriakides, 1991), should completely be discarded from material modeling. Hereafter, we only intend to remove undesirable ratchetting effects, and leave the description of ratchetting for future research. After elimination of uncontrollable artificial ratchetting, physical ratchetting could independently be described for instance by adding internal variables to the scaling relation $S(\delta)$.

As previously stated, the original bounding surface plasticity (BS) model has only two dipoles (i.e., $M = 2$), and therefore has the smallest possible memory size. We propose an improved BS model without changing a single model parameter, simply by increasing the memory size M. The performance of the original and improved BS models will now be compared in two examples.

Figures 4(a) and 4(c) show the stress-strain responses simulated by $M = 2$ and $M = 4$ for a particular set of material constants ($r = s = h_0 = 1$, $\epsilon_i = \sigma_{max}/E$, and $\sigma_{min} = -(\sigma_{max}/2)$, during unsymmetric cycles where σ/σ_{max} varies between 0.8 and 0.3. As shown in Fig. 4(a), the response simulated for $M = 2$ displays a systematic strain accumulation. As previously mentioned, this artificial ratchetting is typical for many bounding surface plasticity models, but is not observed experimentally. When $H > 0$ on the bounding surface, the stress-strain loops drift until they become centered within the bounding surface. However, when $H = 0$ on the bounding surface, the drift persists, and generates infinite strain as in Fig. 4(a). Some models (e.g., Bardet, 1986; Dafalias and Herrmann, 1982) assume purely elastic unloading, and generate even larger ratchetting strain. As shown in Fig. 4(b), when $M = 2$, SM alternates between two triangles during cyclic loading. As shown in Fig. 4(c), when $M = 4$, the stress-strain loops become closed, and the artificial ratchetting is eliminated. As shown in Fig. 4(d), the corresponding SM initially has two dipoles, and then switches between two alternative four-dipole configurations.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

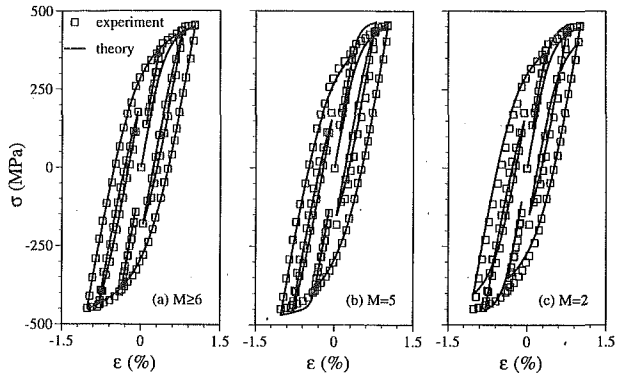


Fig. 5 Experimental stress-strain response of 316 stainless steel (after Chaboche, 1991), and stress-strain responses simulated by (*a*) $M \geq 6$, (*b*) $M = 5$, and (*c*) $M = 2$

Any improved model with $M \geq 4$ would have produced the same simulation as that of Fig. 4(*c*).

Figure 5 shows the comparison between simulated and measured stress-strain responses for a 316 stainless steel subjected to uniaxial cycles of various amplitudes (experiment after Chaboche, 1991). The measured stress-strain loops remain closed. The material constants of the models were calibrated to fit the largest stress-strain loop of Fig. 5(*a*) ($E = 100$ GPa, $\sigma_{\max} = 500$ MPa, and $\sigma_{\min} = -500$ MPa, $h_0 = 7.4$, $r = 2.05$, and $s = 0.58$). The cyclic response was simulated by prescribing the extremities of the strain cycles. As shown in Fig. 5(*a*), the improved model ($M = 6$) accurately simulates the small stress-strain loops within the largest loop. At the beginning SM has two dipoles, and then requires a total of 6 dipoles during the ten following loading reversals, implying that any improved model with $M \geq 6$ would give the same result as that of Fig. 5(*a*). As shown in Fig. 5(*b*), when $M = 5$, the model slightly overestimates the experimental behavior, but displays closed stress-strain loops. This indicates that the scaled memory can be approximated with five dipoles without significantly changing the simulated response. Finally, as shown in Fig. 5(*c*), the original bounding surface model ($M = 2$) describes the largest stress-strain loop, but fails to predict the closure of the smaller loops. This failure to close the smaller loops caused the artificial ratchetting of Fig. 4(*a*).

Six-Dimensional SM Generalization

The one-dimensional SM concept can be generalized to six dimensions in various ways. A particular generalization is now presented for pressure-independent materials such as metals. In six dimensions, the dipoles are generalized into hypersurfaces, called memory surfaces to distinguish them from yield surfaces. The scaled modulus may also be defined in various ways. Hereafter, it is defined as the deviatoric distance between the stress state and memory surfaces. We only intend to illustrate the SM generalization, and do not claim to cover all the subtleties of cyclic metal behavior with a constitutive model having six material constants.

Generalized SM Model. Hereafter, the symbol : denotes the contracted product between two second-order tensors (i.e., $\mathbf{s} : \mathbf{s} = s_{ij}s_{ij}$); the symbol $\| \ \|$ represents their norms (i.e., $\|\mathbf{s}\|^2 = \mathbf{s} : \mathbf{s}$); and the brackets $\langle \rangle$ represent the McCauley symbol:

$$\langle z \rangle = \begin{cases} z & \text{if } z > 0 \\ 0 & \text{if } z \leq 0. \end{cases} \tag{30}$$

As in elastoplasticity, the incremental strain $\mathbf{d}\epsilon$ is the sum of an elastic strain $\mathbf{d}\epsilon^e$ and a plastic strain $\mathbf{d}\epsilon^p$:

$$\mathbf{d}\epsilon = \mathbf{d}\epsilon^e + \mathbf{d}\epsilon^p. \tag{31}$$

$\mathbf{d}\epsilon^e$ may be related to the stress increment $\mathbf{d}\sigma$ by using isotropic elasticity:

$$\mathbf{d}\epsilon^e = \frac{1}{2G}\,\mathbf{d}s + \frac{d\sigma_{kk}}{9B}\,\boldsymbol{\delta}, \tag{32}$$

where B is the elastic bulk modulus, G is the elastic shear modulus, $\boldsymbol{\delta}$ is Kroneker's symbol, and $\mathbf{d}s$ is the increment of deviator stress $\mathbf{s}$:

$$\mathbf{s} = \boldsymbol{\sigma} - \tfrac{1}{3}(\sigma_{kk})\boldsymbol{\delta}. \tag{33}$$

The plastic strain increment is assumed to follow an associative flow rule:

$$\mathbf{d}\epsilon^p = \left\langle \frac{1}{H}\,\mathbf{n} : \mathbf{d}\sigma \right\rangle \mathbf{n} \tag{34}$$

where $\mathbf{n}$ is the unit tensor collinear to the flow (and yield) direction (i.e., $\mathbf{n} : \mathbf{n} = 1$). We assume that the dipole $D_i(c_i, t_i, \delta_i)$ becomes a von Mises memory surface:

$$\|\mathbf{s} - \boldsymbol{\alpha}_i\|^2 - R_i^2 = 0. \tag{35}$$

The deviator tensor $\boldsymbol{\alpha}_i$ represents the position of the surface axis, and R_i is its radius. On the memory surface i, the scaled modulus δ takes the constant value δ_i. In six dimensions, the scaled memory becomes a stack of nested von Mises surfaces. The largest surface is the failure surface while the smallest one is reduced to a stress point. As in the one-dimensional case, the memory surfaces are numbered so that the δ_i's form an increasing series.

The scaled modulus δ at stress $\mathbf{s}$ located between memory surfaces i and $i + 1$ is defined by linear interpolation:

$$\delta = \delta_{i+1} + x(\delta_i - \delta_{i+1}) \tag{36}$$

where δ_i and δ_{i+1} are the scaled moduli on memory surfaces i and $i + 1$, respectively. The scalar x is found by introducing the intermediate surface (see Fig. 6) which passes through the present stress state $\mathbf{s}$ and has the following center $\boldsymbol{\alpha}$ and radius R:

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_{i+1} + x(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}), \quad \text{and}$$

$$R = R_{i+1} + x(R_i - R_{i+1}). \tag{37}$$

Since $\mathbf{s}$ is on the intermediate surface, the scalar x obeys the following quadratic equation:

$$(\mathbf{s} - \boldsymbol{\alpha}_{i+1} - x(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1})) : (\mathbf{s} - \boldsymbol{\alpha}_{i+1} - x(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1})) = (R_{i+1} + x(R_i - R_{i+1}))^2 \tag{38}$$

Therefore,

$$x = \frac{-b - \sqrt{b^2 - ac}}{a}, \tag{39}$$

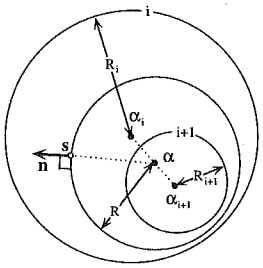


Fig. 6 Representation of scaled modulus δ, stress images, and yield direction; and generation of memory surface at loading reversal

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$\begin{cases} a = (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}) : (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}) - (R_i - R_{i+1})^2 \\ b = -(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}) : (\mathbf{s} - \boldsymbol{\alpha}_{i+1}) - R_{i+1}\,(R_i - R_{i+1}) \\ c = (\mathbf{s} - \boldsymbol{\alpha}_{i+1}) : (\mathbf{s} - \boldsymbol{\alpha}_{i+1}) - R_{i+1})^2. \end{cases} \quad (40)$$

The scalar x is always uniquely defined because surface $i + 1$ is inside surface i (i.e., $a < 0$) and $\mathbf{s}$ is outside surface $i + 1$ (i.e., $c > 0$). The value of x is between 0 and 1. $x = 1$ when $\mathbf{s}$ is on surface i, whereas $x = 0$ when $\mathbf{s}$ is on surface $i + 1$. It can also be shown that the intermediate surface and memory surfaces i and $i + 1$ are always nested, and only intersect when they become tangent, i.e.:

$$\|\boldsymbol{\alpha} - \boldsymbol{\alpha}_{i+1} \leq R - R_{i+1}, \quad \text{and} \quad \|\boldsymbol{\alpha} - \boldsymbol{\alpha}_i\| \leq R_i - R, \quad (41)$$

provided that the surfaces i and $i + 1$ are initially nested (i.e., $\|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}\| \leq R_i - R_{i+1}$). The plastic modulus H in Eq. (34) is defined from δ in Eq. (36) by using the scaling function calibrated in the one-dimensional case. The flow (and yield) direction $\mathbf{n}$ is the unit vector normal to the intermediate surface at stress $\mathbf{s}$:

$$\mathbf{n} = \frac{\mathbf{s} - \boldsymbol{\alpha}}{R}. \quad (42)$$

When $\mathbf{s}$ touches the ith memory surface, $x = 1$ in Eq. (37), and $\mathbf{n}$ becomes the normal to the ith surface. This guarantees that the stress $\mathbf{s}$ crosses the memory surfaces on which $H > 0$ only during loading (i.e., $\mathbf{n} : \mathbf{d}s > 0$). Equation (42) also implies that there is no way even during unloading, for $\mathbf{s}$ to go outside the failure surface on which $H = 0$.

The process defined in Eqs. (36) to (42) also applies at the beginning of loading where the surface $i + 1$ has a zero radius:

$$\delta_{i+1} = 1, \quad \boldsymbol{\alpha}_{i+1} = \mathbf{s}_0 \quad \text{and} \quad R_{i+1} = 0. \quad (43)$$

When an unloading takes place at $\mathbf{s} = \mathbf{s}_U$ between memory surfaces i and $i + 1$ (i.e., $\mathbf{n} : \mathbf{d}s < 0$), the intermediate surface replaces the memory surface $i + 1$, and the reversal stress $\mathbf{s}_U$ is introduced as the last memory surface:

$$\delta_{i+2} = 1, \quad \boldsymbol{\alpha}_{i+2} = \mathbf{s}_u \quad \text{and} \quad R_{i+2} = 0. \quad (44)$$

The total number of memory surfaces becomes $m = i + 2$. After the memory surface has been redefined, Eqs. (37) to (42) can be applied to describe the unloading response.

Approximation of Memory Surfaces. The approximation of scaled memory can be generalized to six dimensions by introducing the intermediate surface p that is interpolated from the surfaces $k - 1$ and $k + 1$ with the scaled modulus of surface k:

$$\boldsymbol{\alpha}_p = \boldsymbol{\alpha}_{k-1} + (\boldsymbol{\alpha}_{k+1} - \boldsymbol{\alpha}_{k-1}) \frac{\delta_k - \delta_{k-1}}{\delta_{k+1} - \delta_{k-1}}, \quad \text{and}$$

$$R_p = R_{k-1} + (R_{k+1} - R_{k-1}) \frac{\delta_k - \delta_{k-1}}{\delta_{k+1} - \delta_{k-1}}. \quad (45)$$

The distance E_k between surfaces p and k is defined as

$$E_k^2 = 2\|\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_p\|^2 + 2(R_p - R_k)^2, \quad (46)$$

which coincides with Eq. (29) in the one-dimensional case. When the number of memory surfaces exceeds the maximum memory size M, the surface k with the minimum E_k is eliminated from the scaled memory.

Simulation of Uniaxial and Torsional Responses. The generalized SM models apply to any multidimensional loadings, and in particular to proportional loadings where $\mathbf{s}$ and all $\boldsymbol{\alpha}_i$'s are collinear. Proportional loadings are commonly applied to laboratory specimens, e.g., uniaxial compression, extension, torsion, simple shear, etc. For proportional loadings, Eq. (39) becomes

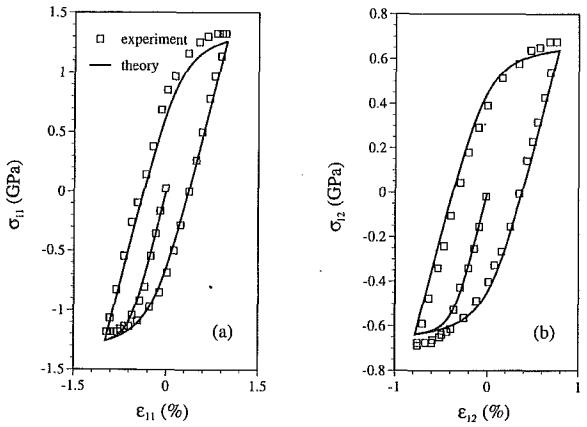


Fig. 7 (*a*) Experimental and simulated responses for push-pull test on a superalloy; and (*b*) experimental and simulated responses for cyclic pure torsion on a superalloy (experimental results after Han and Wack, 1987).

$$x = \frac{\|\mathbf{s} - \boldsymbol{\alpha}_{i+1}\| - R_{i+1}}{R_i - R_{i+1} + \|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}\| \operatorname{sign}((\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{i+1}) : (\mathbf{s} - \boldsymbol{\alpha}_{i+1}))}. \quad (47)$$

When $\boldsymbol{\alpha}_i = 0$, the incremental stress-strain relation in the case of uniaxial push-pull tests ($\sigma_{12} = \sigma_{23} = \sigma_{31} = \sigma_{22} = \sigma_{33} = 0$) becomes

$$d\epsilon_{11} = \frac{d\sigma_{11}}{E} + \frac{2\,d\sigma_{11}}{3H}, \quad (48)$$

and in the case of torsional tests ($\sigma_{23} = \sigma_{13} = \sigma_{11} = \sigma_{22} = \sigma_{33} = 0$),

$$d\epsilon_{12} = \frac{d\sigma_{12}}{2G} + \frac{d\sigma_{12}}{H}. \quad (49)$$

Figure 7 shows the experimental results obtained for a superalloy during push-pull cycles and torsion cycles (Han and Wack, 1987). The amplitudes of the strain cycles are $\Delta\epsilon_{11} = 0.78$ percent and $\Delta\epsilon_{12} = 0.98$ percent for the uniaxial and torsion tests, respectively. The elastic Young's modulus and Poisson ratio are $E = 210.2$ GPa and $\nu = 0.285$. The von Mises strength is $R = 1.32$ GPa. The values of the material constants, $h_0 = 12$, $r = 1.8$, and $s = 1$, have been calibrated by fitting the uniaxial response, and have been used to predict the torsional response. The simulated responses have been obtained through a step-by-step integration of Eqs. (48) and (49). As shown in Fig. 7, the simulation and predictions are in agreement with the experimental results. The generalized SM model, like kinematic hardening, describes closed stress-strain loops for various loading directions, and large-amplitude cycles.

SM was only generalized for pressure-independent plasticity. However, the SM concept can be extended for pressure-dependent materials, e.g., Drucker-Prager and Mohr-Coulomb. SM is also versatile enough to be inserted into existing bounding surface models for describing the variation of plastic modulus during arbitrary cyclic loadings.

Conclusion

The Scaled Memory (SM) concept provides a geometrical representation of cyclic stress-strain responses. SM scales the nonlinear variation of tangential modulus during monotonic loadings into a piecewise linear distribution, which can easily be modified to enforce specific features of materials behaviors

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

during cyclic loadings. We have illustrated the usefulness of SM by eliminating artificial ratchetting in bounding surface plasticity, without introducing a single material constant. We have also illustrated the six-dimensional generalization of SM by constructing a constitutive model for pressure-independent materials. Future research is needed to apply SM to pressure-dependent materials, and materials with internal variables.

Acknowledgment

The financial support of the National Science Foundation is acknowledged (grant MSM8657999). The author thanks M. Harris for proofreading the manuscript.

References

Abu-Lebdeh, T. M., and Voyiadjis, G. Z., 1993, "Plasticity-damage model for concrete under cyclic multiaxial loading," *Journal of Engineering Mechanics,* ASCE, Vol. 119, No. 7, pp. 1465–1484.

Bardet, J. P., 1986, "A bounding surface model for sands," *Journal of Engineering Mechanics,* ASCE, Vol. 112, No. 11, pp. 1198–1216.

Bardet, J. P., 1990, "Hypoplastic model for sands," *Journal of Engineering Mechanics,* ASCE, Vol. 116, No. 9, pp. 1973–1994.

Bardet, J. P., 1995, "Scaled memory model for cyclic behavior of soils," *Journal of Geotechnical Engineering,* ASCE, Vol. 121, No. 11, pp. 766–775.

Chaboche, J. L., 1986, "Time-independent constitutive theories for cyclic plasticity," *International Journal of Plasticity,* Vol. 2, pp. 149–188.

Chaboche, J. L., 1991, "On some modifications of kinematic hardening to improve the description of ratchetting effects," *International Journal of Plasticity,* Vol. 7, pp. 661–678.

Chu, C. C., 1984, "A three-dimensional model of anisotropic hardening in metals and its application to the analysis of sheet metal formability," *Journal of Mechanics and Physics of Solids,* Vol. 32, No. 3, pp. 197–212.

Chu, C. C., 1987, "The analysis of multiaxial cyclic problems with an anisotropic hardening model," *International Journal of Solids and Structures,* Vol. 23, No. 5, pp. 569–579.

Dafalias, Y. F., and Herrmann, L. R., 1982, "Bounding surface formulation of soil plasticity," *Soil Mechanics-Transient and cyclic loads,* G. Pande and O. C. Zienkiewicz, eds., John Wiley and Sons, Inc., London, U.K., pp. 253–282.

Dafalias, Y. F., and Popov, E., 1975, "A model of nonlinearity hardening materials for complex loading," *Acta Mechanica,* Vol. 21, No. 3, pp. 173–192.

Han, S., and Wack, B., 1987, "Discrete memory description of strain hardening and softening with application to stainless steel and superalloy," *Res Mechanica,* Vol. 20, pp. 315–342.

Hashiguchi, K., 1980, "Constitutive Equation of Elastoplastic Materials With Elastic-Plastic Transition," ASME JOURNAL OF APPLIED MECHANICS, Vol. 102, pp. 266–272.

Hashiguchi, K., 1993, "Mechanical requirements and structures of cyclic plasticity models," *International Journal of Plasticity,* Vol. 2, pp. 721–748.

Hassan, T., and Kyriakides, S., 1991, "Ratchetting in cyclic plasticity, Part I: Uniaxial behavior," *International Journal of Plasticity,* Vol. 8, pp. 91–116.

Hill, R., 1950, *The mathematical theory of plasticity,* Oxford University Press, Oxford, U.K.

Iwan, W. D., 1967, "On a class of models for the yielding behavior of continuous and composite systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 612–617.

Kondner, R. L., 1963, "Hyperbolic stress-strain response: cohesive soil," *Journal of the Soil Mechanics and Foundations Division,* ASCE, Vol. 89, No. 1, pp. 115–143.

Krieg, R. D., 1975, "A practical two surface plasticity theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 641–646.

Masing, G., 1926, "Eigenspannungen und Verfestigung beim Messing," *Proceedings of the Second International Congress of Applied Mechanics,* pp. 332–335.

Mróz, Z., 1967, "On the description of anisotropic workhardening," *Journal of Mechanics and Physics of Solids,* Vol. 15, pp. 163–175.

Mróz, Z., 1984, "Hardening and degradation rules for metals under monotonic and cyclic loading," ASME *Journal of Engineering Materials Technology,* Vol. 105, pp. 113–118.

Mróz, Z., and Lind, N. C., 1975, "Simplified theories of cyclic plasticity," *Acta Mechanica,* pp. 131–152.

Mróz, Z., and Norris, V. A., 1982, "Elastoplastic and viscoplastic constitutive models for soils with application to cyclic loading," *Soil Mechanics-Transient and cyclic loads,* G. Pande and O. C. Zienkiewicz, eds., John Wiley and Sons, Inc., London, U.K., pp. 173–217.

Valanis, K. C., 1971, "A theory of viscoplasticity without a yield surface, Part I. General theory," *Archives of Mechanics,* Vol. 53, pp. 517–533.

Wack, B., Terriez, J. M., and Guelin, P., 1983, "A hereditary type, discrete memory, constitutive equation with application to simple geometries," *Acta Mechanica,* Vol. 50, pp. 9–37.

Wang, Z. L., Dafalias, Y. F., and Shen, C., 1990, "Bounding surface hypoplasticity model for sand," *Journal of Engineering Mechanics,* ASCE, Vol. 116, No. 5, pp. 983–1001.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Simulation of Multivariate Gaussian Fields Conditioned by Realizations of the Fields and Their Derivatives

Y. J. Ren
Research Assistant.

I. Elishakoff
Professor,
Mem. ASME.

Department of Mechanical Engineering,
Florida Atlantic University,
Boca Raton, FL 33431-0991

M. Shinozuka
Fred Chapman Chair Professor,
Department of Civil Engineering,
University of Southern California,
Los Angeles, CA 90089-0242
Fellow ASME

This paper investigates conditional simulation technique of multivariate Gaussian random fields by stochastic interpolation technique. For the first time in the literature a situation is studied when the random fields are conditioned not only by a set of realizations of the fields, but also by a set of realizations of their derivatives. The kriging estimate of multivariate Gaussian field is proposed, which takes into account both the random field as well as its derivative. Special conditions are imposed on the kriging estimate to determine the kriging weights. Basic formulation for simulation of conditioned multivariate random fields is established. As a particular case of uncorrelated components of multivariate field without realizations of the derivative of the random field, the present formulation includes that of univariate field given by Hoshiya. Examples of a univariate field and a three component field are elucidated and some numerical results are discussed. It is concluded that the information on the derivatives may significantly alter the results of the conditional simulation.

Introduction

Recently, conditional simulation of random processes and fields has been studied due to its applications to urban earthquake monitoring. The conditional nature of the simulation stems from the fact that the realizations of the random processes or fields at some locations have been recorded. One needs to simulate the random field in other locations when recorded information is given.

So far two conditional simulation techniques have been dealt by investigators. Vanmarcke and Fenton (1991) applied the spectral representation method to expand random processes into the Fourier series. The conditional simulation of Fourier coefficients was implemented by the kriging technique. Kameda and Morikawa (1992, 1994) established an analytical framework by the spectral representation method and by expanding conditioned random processes into the Fourier series and derived expressions of joint probability density functions of the Fourier coefficients. They also obtained expressions of conditional expectations and variances of the conditioned random processes and considered their first-passage probabilities. Hoshiya (1994) and Hoshiya and Maruyama (1994) applied the kriging technique to get kriging estimates of random fields and then to simulate the error between the actual random field and its kriging estimate. (For some other applications of the kriging technique, the reader can consult with the monograph by Journel and Huijbregts (1978), papers by Krige (1951), Pourier and Tinawi (1991), Jeulin and Renard (1992), and Trochu (1993).) The simulation of a conditioned random field was accomplished by summing the kriging estimate and the error. Ishii and Suzuki (1989) also utilized the kriging technique to estimate the spatial variation of soil parameters such as Young's modules, which was combined with the stochastic version of the finite element method to analyze structural reliability.

The existing conditional simulation techniques have been devoted to univariate random processes and random fields, which are conditioned by realizations of the field itself. The problem can be briefly described as follows: Let $G(x, t)$ be a Gaussian random field and $g(x_i, t)$ be recorded realizations of $G(x, t)$ at some locations x_i $(i = 1, 2, \ldots, N)$; one is required to simulate the random field at other (unrecorded) locations x_r, namely $G[x_r, t|G(x_i, t) = g(x_i, t), i = 1, 2, \ldots, N]$ $(r = 1, 2, \ldots, R)$. This study investigates the conditional simulation technique for a *wider* class, namely for multivariate random fields, which are conditioned not only by realizations of the random field but also by realizations of its derivatives. This is motivated by the fact that in practical situations not only displacements but also its derivatives (slopes, strains) may be simultaneously measured.

In the present study, the kriging estimate of a multivariate random field is proposed, which consists of the random field itself as well as the derivatives of the random field. Instead of the conventional requirement of minimizing the variance of the error, which is the difference between the actual field and its kriging estimate, special conditions are proposed to guarantee the kriging estimate to be an unbiased linear estimator. These conditions appear to be natural to be introduced for the kriging estimate consisting of both the random field and its derivative. Basic formulation for conditional simulation of multivariate random fields is then developed. Important properties of the error vector are discussed. These properties guarantee that the conditional simulation of the multivariate field can be implemented by computing its kriging estimate and simulating the error vector *separately*, and summing them subsequently to simulate the conditional field. It is also shown that the present formulation includes that of univariate random fields proposed by Hoshiya, when the components of a multivariate field are uncorrelated and no realizations of derivatives are considered. Two examples and some numerical results are presented to illustrate the effectiveness of the proposed conditional simulation.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Aug. 22, 1994; final revision, July 5, 1995. Technical Editor: L. T. Wheeler.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Basic Formulation

Assume that $\mathbf{G}(x) = [G_1(x), G_2(x), \ldots, G_n(x)]^T$ is a homogeneous n-variate Gaussian random vector field with zero-mean and cross-variance matrix $C[G_k(x_i), G_l(x_j)] = E[G_k(x_i) \times G_l(x_j)]$ $(k, l = 1, 2, \ldots, n)$. Let $\mathbf{G}_l = [G_l(x_1), G_l(x_2), \ldots, G_l(x_N)]^T$ be a vector of random variables of the lth component of $\mathbf{G}(x)$ at locations x_i $(i = 1, 2, \ldots, N)$ and $\mathbf{G}'_l = [G'_l(x_1), G'_l(x_2), \ldots, G'_l(x_M)]^T$ be a vector of random variables of the lth component of the derivative field $\mathbf{G}'(x) = d\mathbf{G}(x)/dx$ at locations x_j $(j = 1, 2, \ldots, M)$. Let $\mathbf{g}_l$ be a realization of the vector $\mathbf{G}_l$ and $\mathbf{g}'_l$ be a realization of $\mathbf{G}'_l$. We approximate the actual field $\mathbf{G}(x)$ by its simulated counterpart $\mathbf{G}^{(s)}(x)$. In component form we have

$$G_l^{(s)}(x) = G_l^{(k)}(x) + [G_l(x) - G_l^{(k)}(x)]$$

$$= G_l^{(k)}(x) + \epsilon_l(x), \quad l = 1, 2, \ldots, n \qquad (1)$$

where the following expression is postulated for $G_l^{(k)}(x)$:

$$G_l^{(k)}(x) = \sum_{m=1}^{n} \mathbf{G}_m^T \lambda_{ml}(x) + \sum_{m=1}^{n} \mathbf{G}_m'^T \boldsymbol{\mu}_{ml}(x). \qquad (2)$$

Equation (2) is proposed here as the kriging estimate of the lth component $G_l(x)$ of the multivariate random field $\mathbf{G}(x)$, and $\lambda_{ml}(x) = [\lambda_{ml1}(x), \lambda_{ml2}(x), \ldots, \lambda_{mlN}(x)]^T$ and $\boldsymbol{\mu}_{ml}(x) = [\mu_{ml1}(x), \mu_{ml2}(x), \ldots, \mu_{mlM}(x)]^T$ are vectors of kriging weight functions. Equation (2) takes into account both the random field and its derivative for the stochastic interpolation of the random field. It represents a generalization of existing kriging estimate proposed for univariate random fields (Hoshiya 1994), which did not include information on realizations of the field's derivatives. It will be shown that Eq. (2) reduces to kriging estimate of univariate fields for the particular case when the components of the field are mutually uncorrelated and the derivative of the random field is not recorded. In Eq. (1), $\epsilon_l(x)$ is the error between the exact field $G_l(x)$ and its kriging estimate $G_l^{(k)}(x)$

$$\epsilon_l(x) = G_l(x) - \sum_{m=1}^{n} \mathbf{G}_m^T \lambda_{ml}(x) - \sum_{m=1}^{n} \mathbf{G}_m'^T \boldsymbol{\mu}_{ml}(x). \qquad (3)$$

In the conventional kriging technique as well as in the conventional regression method, it is required that the variance of the error $\epsilon_l(x)$ attain a minimum. Note that the kriging estimate in Eq. (3) consists of both the random variables of the field and its derivative at specified locations. The expression in Eq. (3), in fact, forms an interpolation of random variables. It is instructive to recall some facts from the interpolation theory of deterministic functions: the interpolating coefficients are determined based on the conditions that at locations where the function is specified, the interpolated function must be equal to the specified values of the function; moreover, at locations where the derivative of the function is specified, the derivative of the interpolated function must be equal to the specified values of the derivative. Following this line of thought, we impose *special* conditions on the kriging estimate $G_l^{(k)}(x)$ instead of requiring the variance of the error $\epsilon_l(x)$ to attain a minimum. Generalizing the idea used for interpolation of a deterministic function, we require simultaneously that, in the *probabilistic sense,* the kriging estimate must be *equivalent* to the actual field where the realizations of the field are measured, and the derivative of the kriging estimate must be *equivalent* to the derivative of the actual field where the realizations of the derivative of the field are measured. Explicitly, the following four conditions should be satisfied:

(1)

$$E[\mathbf{G}_k G_l(x)] = E[\mathbf{G}_k G_l^{(k)}(x)], \quad k = 1, 2, \ldots, n \qquad (4)$$

Substitution of Eq. (2) yields

$$\sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = C[\mathbf{G}_k, G_l(x)]$$

$$k = 1, 2, \ldots, n \qquad (5)$$

where $\lambda_{ml} = \lambda_{ml}(x)$ and $\boldsymbol{\mu}_{ml} = \boldsymbol{\mu}_{ml}(x)$ for simplicity.

(2)

$$E[\mathbf{G}_k G'_l(x)] = E[\mathbf{G}_k G_l'^{(k)}(x)], \quad k = 1, 2, \ldots, n \qquad (6)$$

From Eq. (2), we have

$$G_l'^{(k)}(x) = \sum_{m=1}^{n} \mathbf{G}_m^T \lambda'_{ml}(x) + \sum_{m=1}^{n} \mathbf{G}_m'^T \boldsymbol{\mu}'_{ml}(x) \qquad (7)$$

Hence

$$\sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m^T]\lambda'_{ml} + \sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m'^T]\boldsymbol{\mu}'_{ml} = C[\mathbf{G}_k, G'_l(x)]$$

$$k = 1, 2, \ldots, n \qquad (8)$$

Integrating Eq. (8) once yields

$$\sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml}$$

$$= C[\mathbf{G}_k, G_l(x)] + \mathbf{A}_l \quad k = 1, 2, \ldots, n \qquad (9)$$

where $\mathbf{A}_l$ = vector of constants. Comparison of Eq. (5) and Eq. (9) results in $\mathbf{A}_l = 0$.

(3)

$$E[\mathbf{G}'_k G_l(x)] = E[\mathbf{G}'_k G_l^{(k)}(x)], \quad k = 1, 2, \ldots, n \qquad (10)$$

Substitution of Eq. (2) yields

$$\sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = C[\mathbf{G}'_k, G_l(x)]$$

$$k = 1, 2, \ldots, n \qquad (11)$$

(4)

$$E[\mathbf{G}'_k G'_l(x)] = E[\mathbf{G}'_k G_l'^{(k)}(x)], \quad k = 1, 2, \ldots, n \qquad (12)$$

Substitution of Eq. (2) and Eq. (7) gives

$$\sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m^T]\lambda'_{ml} + \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m'^T]\boldsymbol{\mu}'_{ml} = C[\mathbf{G}'_k, G'_l(x)]$$

$$k = 1, 2, \ldots, n. \qquad (13)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Integrating Eq. (13) once results in

$$\sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = C[\mathbf{G}'_k, G_l(x)] + \mathbf{B}_l \quad k = 1, 2, \ldots, n \quad (14)$$

where $\mathbf{B}_l$ = vector of constants. Comparison of Eq. (11) and (14) also yields $\mathbf{B}_l = 0$.

It is remarkable that the first and third conditions are equivalent to the requirement of minimizing the variance of the error $\epsilon_l(x)$, whereas the second and fourth conditions are equivalent to the requirement of minimizing the variance Var $[\epsilon'_l(x)]$ of the error's derivative $\epsilon'_l(x)$. Whereas the former requirement, namely to minimize Var $[\epsilon_l(x)]$, is obvious, the latter one, namely to minimize Var $[\epsilon'_l(x)]$, appears to be less transparent at the first glance.

Note that once we obtain $\mathbf{B}_l = 0$, Eq. (14) coincides with Eq. (11). In perfect analogy, since $\mathbf{A}_l = 0$, Eq. (9) coincides with Eq. (5). Indeed, one may show that the condition Var $[\epsilon'_l(x)]$ = min is a consequence of the condition Var $[\epsilon_l(x)]$ = min. One may, however, visualize the situation in which these conditions are independent. In present formulation, four requirements reduce to two independent conditions. Combination of requirements (1)–(4) yields the following linear equations for the kriging weight functions $\lambda_{ml}(x)$ and $\boldsymbol{\mu}_{ml}(x)$:

$$\sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = C[G_l(x), \mathbf{G}_k]$$

$$\sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m^T]\lambda_{ml} + \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = C[G_l(x), \mathbf{G}'_k]$$

$$k = 1, 2, \ldots, n. \quad (15)$$

Equation (15) consists of $n \times (N + M)$ equations, from which $n \times (N + M)$ unknown kriging weights λ_{ml} and $\boldsymbol{\mu}_{ml}$ for the component $G_l(x)$ can be determined.

It is appropriate to mention that the classical regression method in statistics can be generalized in a way similar to present formulation, in order to estimate multivariate random fields. Both the generalized classical regression method and generalized kriging method will then coincide with each other provided that the same set of requirements are imposed, and Gaussian fields are considered.

The error $\epsilon_l(x)$ possesses the following properties:

(i) The mean of the error vector vanishes,

$$E[\epsilon_l(x)] = E[G_l(x) - \sum_{m=1}^{n} \mathbf{G}_m^T\lambda_{ml} - \sum_{m=1}^{n} \mathbf{G}_m'^T\boldsymbol{\mu}_{ml}]$$

$$= E[G_l(x)] - \sum_{m=1}^{n} E[\mathbf{G}_m^T]\lambda_{ml} - \sum_{m=1}^{n} E[\mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = 0. \quad (16)$$

(ii) The error vector and the random vector $\mathbf{G}_k$ are uncorrelated,

$$E[\epsilon_l(x)\mathbf{G}_k] = E[\mathbf{G}_k[G_l(x) - \sum_{m=1}^{n} \mathbf{G}_m^T\lambda_{ml} - \sum_{m=1}^{n} \mathbf{G}_m'^T\boldsymbol{\mu}_{ml}]\}$$

$$= C[\mathbf{G}_k, G_l(x)] - \sum_{m=1}^{n} C[\mathbf{G}_k, \mathbf{G}_m^T]\lambda_{ml} - \sum_{m=1}^{n} \mathbf{C}[\mathbf{G}_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = 0$$

$$k = 1, 2, \ldots, n. \quad (17)$$

(iii) The error vector and the random vector $\mathbf{G}'_k$ are uncorrelated,

$$E[\epsilon_l(x)\mathbf{G}'_k] = E\{\mathbf{G}'_k[G_l(x) - \sum_{m=1}^{n} \mathbf{G}_m^T\lambda_{ml} - \sum_{m=1}^{n} \mathbf{G}_m'^T\boldsymbol{\mu}_{ml}]\}$$

$$= C[\mathbf{G}'_k, G_l(x)] - \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m^T]\lambda_{ml} - \sum_{m=1}^{n} C[\mathbf{G}'_k, \mathbf{G}_m'^T]\boldsymbol{\mu}_{ml} = 0$$

$$k = 1, 2, \ldots, n. \quad (18)$$

(iv) At any recorded location x_i, Eq. (15) becomes an identity by simply letting $\lambda_{mli} = 1$ and taking other kriging weights equal to zero, i.e., $\lambda_{mlj} = 0$ ($j \neq i$ or $m \neq l$) and every $\mu_{mkj} = 0$. We then have

$$\epsilon_l(x_i) = G_l(x_i) - \sum_{m=1}^{n} \mathbf{G}_m^T\lambda_{ml} - \sum_{m=1}^{n} \mathbf{G}_m'^T\boldsymbol{\mu}_{ml} = G_l(x_i) - G_l(x_i) = 0. \quad (19)$$

Subsequently,

$$E[\epsilon_l(x)\epsilon_k(x_i)] = 0. \quad (20)$$

(v) The error vector and the kriging estimate are uncorrelated,

$$E[\epsilon_l(x)G_k^{(k)}(x)] = E\{[\sum_{s=1}^{n} \lambda_{sk}^T\mathbf{G}_s + \sum_{s=1}^{n} \mu_{sk}^T\mathbf{G}'_s][G_l(x) - \sum_{m=1}^{n} \mathbf{G}_m^T\lambda_{ml} - \sum_{m=1}^{n} \mathbf{G}_m'^T\boldsymbol{\mu}_{ml}]\} = \sum_{s=1}^{n} \lambda_{sk}^T\{C[\mathbf{G}_s, G_l(x)] - \sum_{m=1}^{n} C[\mathbf{G}_s, \mathbf{G}_m^T]\lambda_{ml} - \sum_{m=1}^{n} C[\mathbf{G}_s, \mathbf{G}'_m]\boldsymbol{\mu}_{ml}\} + \sum_{s=1}^{n} \boldsymbol{\mu}_{sk}^T\{C[\mathbf{G}'_s, G_l(x)] - \sum_{m=1}^{n} C[\mathbf{G}'_s, \mathbf{G}_m]\lambda_{ml} - \sum_{m=1}^{n} C[\mathbf{G}'_s, \mathbf{G}'_m]\boldsymbol{\mu}_{ml}\} = 0 \quad k = 1, 2, \ldots, n. \quad (21)$$

(vi) Correlation matrix of the error vector is obtainable as

$$E[\epsilon_l(x)\epsilon_k(x)] = E\{\epsilon_l(x)[G_k(x) - G_k^{(k)}(x)]\} = E[\epsilon_l(x)G_k(x)] = C[G_l(x), G_k(x)] - \sum_{m=1}^{n} \lambda_{ml}^T C[\mathbf{G}_m, G_k(x)] - \sum_{m=1}^{n} \boldsymbol{\mu}_{ml}^T C[\mathbf{G}'_m, G_k(x)]. \quad (22)$$

In particular, the variance of the error reads

$$\text{Var}\,[\epsilon_l(x)] = \text{Var}\,[G_l(x)] - \sum_{m=1}^{n} \lambda_{ml}^T C[\mathbf{G}_m, G_l(x)] - \sum_{m=1}^{n} \boldsymbol{\mu}_{ml}^T C[\mathbf{G}'_m, G_l(x)]. \quad (23)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

We arrive at the following:

Theorem. The variance of the error $\epsilon_l(x)$ reaches its minimum, given in Eq. (23), if and only if the weights λ_{ml} and $\boldsymbol{\mu}_{ml}$ satisfy Eq. (15).

Proof: Assume that λ^*_{ml} and $\boldsymbol{\mu}^*_{ml}$ form an arbitrary set of kriging weights and $\epsilon^*_l(x)$ is the associated error. Denote $\lambda^*_{ml} = \lambda_{ml} + \delta\lambda_{ml}$ and $\boldsymbol{\mu}^*_{ml} = \boldsymbol{\mu}_{ml} + \delta\boldsymbol{\mu}_{ml}$, where λ_{ml} and $\boldsymbol{\mu}_{ml}$ are the kriging weights satisfying Eq. (11), and $\delta\lambda_{ml}$ and $\delta\mu_{ml}$ are deviations. From Eq. (3), we have

$$\begin{aligned}
\mathrm{Var}\,[\epsilon^*_l(x)] &= \mathrm{Var}\,[G_l(x)] - 2\sum_{k=1}^{n}\lambda^{*T}_{kl}C[\mathbf{G}_k, G_l(x)] - 2\sum_{k=1}^{n}\boldsymbol{\mu}^{*T}_{kl}C[\mathbf{G}'_k, G_l(x)] \\
&\quad + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\lambda^{*T}_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\boldsymbol{\mu}^*_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\lambda^{*T}_{kl}C[\mathbf{G}_k, \mathbf{G}^T_m]\lambda^*_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\boldsymbol{\mu}^{*T}_{kl}C[\mathbf{G}'_k, \mathbf{G}'^{T}_m]\boldsymbol{\mu}^*_{ml} \\
&= \mathrm{Var}\,[G_l(x)] - 2\sum_{k=1}^{n}\lambda^T_{kl}C[\mathbf{G}_k, G_l(x)] - 2\sum_{k=1}^{n}\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, G_l(x)] \\
&\quad + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\boldsymbol{\mu}_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}^T_m]\lambda_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, \mathbf{G}'^{T}_m]\boldsymbol{\mu}_{ml} \\
&\quad - 2\sum_{k=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, G_l(x)] - 2\sum_{k=1}^{n}\delta\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, G_l(x)] + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\boldsymbol{\mu}_{ml} \\
&\quad + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, \mathbf{G}^T_m]\lambda_{ml} + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}^T_m]\lambda_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k\,\mathbf{G}'^{T}_m]\boldsymbol{\mu}_{ml} \\
&\quad + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, G'^{T}_m]\delta\boldsymbol{\mu}_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\delta\lambda_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, \mathbf{G}'^{T}_m]\delta\boldsymbol{\mu}_{ml}. \qquad (24)
\end{aligned}$$

Bearing in mind Eq. (15), Eq. (24) becomes

$$\begin{aligned}
\mathrm{Var}\,[\epsilon^*_l(x)] &= \mathrm{Var}\,[G(x)] - \sum_{k=1}^{n}\lambda^T_{kl}C[\mathbf{G}_k, G_l(x)] - \sum_{k=1}^{n}\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, G_l(x)] \\
&\quad + 2\sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\delta\boldsymbol{\mu}_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\delta\lambda^T_{kl}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\delta\lambda_{ml} + \sum_{k=1}^{n}\sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{kl}C[\mathbf{G}'_k, \mathbf{G}'^{T}_m]\delta\boldsymbol{\mu}_{ml} \\
&= \mathrm{Var}\,[\epsilon_l(x)] + E[(\sum_{m=1}^{n}\delta\lambda^T_{ml}\mathbf{G}_m + \sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{ml}\mathbf{G}'_m)^2] \geq \mathrm{Var}\,[\epsilon_l(x)]. \qquad (25)
\end{aligned}$$

Necessity: Let $\mathrm{Var}\,[\epsilon^*_l(x)] = \mathrm{Var}\,[\epsilon_l(x)]$, then by virtue of Eq. (25), $\sum_{m=1}^{n}\delta\lambda^T_{ml}\mathbf{G}_m + \sum_{m=1}^{n}\delta\boldsymbol{\mu}^T_{ml}\mathbf{G}'_m = 0$. We multiply it by $\mathbf{G}_k$ and $\mathbf{G}'_k$ ($k = 1, 2, \ldots, n$), respectively, to obtain

$$\begin{aligned}
&\sum_{m=1}^{n}C[\mathbf{G}_k, \mathbf{G}^T_m]\delta\lambda_{ml} + \sum_{m=1}^{n}C[\mathbf{G}_k, \mathbf{G}'^{T}_m]\delta\mu_{ml} = 0 \\
&\sum_{m=1}^{n}C[\mathbf{G}'_k, \mathbf{G}^T_m]\delta\lambda_{ml} + \sum_{m=1}^{n}C[\mathbf{G}'_k, \mathbf{G}'^{T}_m]\delta\mu_{ml} = 0 \\
&\qquad k = 1, 2, \ldots, n. \qquad (26)
\end{aligned}$$

The solution of Eq. (26) is $\delta\lambda_{ml} = \delta\boldsymbol{\mu}_{ml} = 0$, provided the cross-covariance matrix of $\mathbf{G}(x)$ is nonsingular. Hence, $\lambda^*_{ml} = \lambda_{ml}$ and $\boldsymbol{\mu}^*_{ml} = \boldsymbol{\mu}_{ml}$.

Sufficiency: Let $\lambda^*_{ml} = \lambda_{ml}$ and $\boldsymbol{\mu}^*_{ml} = \boldsymbol{\mu}_{ml}$. Then the deviations vanish identically, $\delta\lambda_{ml} = \delta\boldsymbol{\mu}_{ml} = 0$. Immediately, this results in $\mathrm{Var}\,[\epsilon^*_l(x)] = \mathrm{Var}\,[\epsilon_l(x)]$. The proof is completed.

Properties (i) – (vi) state that the error vector $\epsilon(x_r) = [\epsilon_1(x_r), \epsilon_2(x_r), \ldots, \epsilon_n(x_r)]^T$ of $\mathbf{G}(x)$ at any location x is independent from the random vector $\mathbf{G}_l$ and the random derivative vector $\mathbf{G}'_l$ ($l = 1, 2, \ldots, n$), which represent the random field $\mathbf{G}(x)$ at recorded locations x_i ($i = 1, 2, \ldots, N$), and its derivative $\mathbf{G}'(x)$ at recorded locations x_j ($j = 1, 2, \ldots, M$). The error vector $\epsilon(x)$ is uncorrelated with error vectors $\epsilon(x_i)$ at recorded locations x_i ($i = 1, 2, \ldots, N$). In addition, it is also uncorrelated with the kriging estimate $G^{(k)}_l(x)$ ($l = 1, 2, \ldots, n$) of the random field at the same location x. However, different error components at unrecorded location x are correlated with each other. These important properties of $\epsilon(x)$ in Eqs. (16) – (23) guarantee that the error vector $\epsilon(x)$ can be simulated separately from the kriging estimate. Hence, to simulate Gaussian vector random field $\mathbf{G}(x)$ at an unrecorded location x under the condition of known realizations $\mathbf{g}_l$ and $\mathbf{g}'_l$ ($l = 1, 2, \ldots, n$), we can calculate the kriging estimate of each component $g^{(k)}_l(x)$ ($l = 1, 2, \ldots, n$) and simulate its error vector $\epsilon(x)$ separately, and subsequently formulate their sum. The lth component of the conditioned multivariate random field $\mathbf{G}(x)$ is given by

$$G_l(x) = \sum_{m=1}^{n}\lambda^T_{ml}\mathbf{g}_m + \sum_{m=1}^{n}\boldsymbol{\mu}^T_{ml}\mathbf{g}'_m + \epsilon_l(x)$$

$$l = 1, 2, \ldots, n. \qquad (27)$$

The error vector $\epsilon(x)$ is an n-component vector random field with zero-mean and covariance matrix given Eq. (22). Its simulation techniques are widely available. If the multivariate random field needs to be simulated at several unrecorded locations x_r ($r = 1, 2, \ldots, R$), a location-by-location recursive procedure can be applied, in analogy to the univariate random field case (Hoshiya, 1993).

Multivariate Field With Uncorrelated Components

In the particular case that components of the multivariate random field $\mathbf{G}(x)$ are mutually uncorrelated, namely $C[G_k(x_i), G_m(x_j)] = 0$ for $k \neq m$, Eq. (15) reduces to

$$\begin{bmatrix} C[\mathbf{G}_k, \mathbf{G}^T_k] & C[\mathbf{G}_k, \mathbf{G}'^{T}_k] \\ C[\mathbf{G}'_k, \mathbf{G}^T_k] & C[\mathbf{G}'_k, \mathbf{G}'^{T}_k] \end{bmatrix}\begin{Bmatrix} \lambda_{kl} \\ \boldsymbol{\mu}_{kl} \end{Bmatrix} = \begin{Bmatrix} C[G_l(x), \mathbf{G}_k] \\ C[G_l(x), \mathbf{G}'_k] \end{Bmatrix}$$

$$k = 1, 2, \ldots, n. \qquad (28)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

If $k \neq l$, Eq. (28) becomes

$$\begin{bmatrix} C[\mathbf{G}_k, \mathbf{G}_k^T] & C[\mathbf{G}_k, \mathbf{G}_k'^T] \\ C[\mathbf{G}_k', \mathbf{G}_k^T] & C[\mathbf{G}_k', \mathbf{G}_k'^T] \end{bmatrix} \begin{Bmatrix} \lambda_{kl} \\ \boldsymbol{\mu}_{kl} \end{Bmatrix} = \{0\}$$

$$k = 1, 2, \ldots, n; \quad k \neq l. \quad (29)$$

Due to the non-singularity of the auto-covariance matrix $C[G_k(x_i), G_k(x_j)]$ of the k-th component $G_k(x)$, only a trivial solution exists for Eq. (29), namely

$$\lambda_{kl} = 0, \quad \boldsymbol{\mu}_{kl} = 0 \quad (k = 1, 2, \ldots, n; k \neq l). \quad (30)$$

For $k = l$, Eq. (28) becomes

$$\begin{bmatrix} C[\mathbf{G}_l, \mathbf{G}_l^T] & C[\mathbf{G}_l, \mathbf{G}_l'^T] \\ C[\mathbf{G}_l', \mathbf{G}_l^T] & C[\mathbf{G}_l', \mathbf{G}_l'^T] \end{bmatrix} \begin{Bmatrix} \lambda_l \\ \boldsymbol{\mu}_l \end{Bmatrix} = \begin{Bmatrix} C[G_l(x), \mathbf{G}_l] \\ C[G_l(x), \mathbf{G}_l'] \end{Bmatrix} \quad (31)$$

where $\lambda_l = \lambda_{ll}$ and $\boldsymbol{\mu}_l = \boldsymbol{\mu}_{ll}$ for simplicity. Equation (31) has a unique solution unless the covariance matrix $C[G_l(x_i), G_l(x_j)]$ of the lth component $G_l(x)$ is singular. The latter case takes place, for example, when the correlation coefficient between $G_l(x_i)$ and $G_l(x_j)$ equals unity in its absolute value. Equation (2) and Eq. (23) then become, respectively,

$$G_l^{(k)}(x) = \mathbf{G}_l^T \lambda_l + \mathbf{G}_l'^T \boldsymbol{\mu}_l \quad (32)$$

$$\text{Var}\,[\epsilon_l(x)] = \text{Var}\,[G_l(x)] - \lambda_l^T C[\mathbf{G}_l, G_l(x)] - \boldsymbol{\mu}_l^T C[\mathbf{G}_l'^T, G_l(x)] \quad (33)$$

The different error components $\epsilon_l(x)$ and $\epsilon_k(x)$ for $k \neq l$ are uncorrelated, namely $E[\epsilon_l(x)\epsilon_k(x)] = 0$. It is seen that in the case when components of a multivariate random field $\mathbf{G}(x)$ are uncorrelated, each component $G_l(x)$ of $\mathbf{G}(x)$ can be simulated separately. In fact, Eqs. (31)–(33) consist of collection of uncoupled expressions for conditional simulation of the univariate random field under realizations of the field and realizations of its derivative. Furthermore, if the realizations of the derivative field are not specified or are not considered, we simply take all $\boldsymbol{\mu}_l = 0$ to obtain from Eqs. (31)–(33)

$$C[\mathbf{G}_l, \mathbf{G}_l^T]\lambda_l = C[G_l(x), \mathbf{G}_l] \quad (34)$$

$$G_l^{(k)}(x) = \mathbf{G}_l^T \lambda_l \quad (35)$$

$$\text{Var}\,[\epsilon_l(x)] = \text{Var}\,[G_l(x)] - \lambda_l^T C[\mathbf{G}_l, G_l(x)]. \quad (36)$$

The above three equations are identical to those proposed by Hoshiya (1994) for a univariate random field under realizations of the random field.

To gain physical insight, we apply the general simulation technique proposed in previous sections to determine the conditional means and variances of the conditioned field in two example problems.

Example 1

First, consider a one-dimensional spatially random field $G(x)$ with the following unconditional correlation function

$$C_{GG}(x_1, x_2) = \sigma_0^2 \exp[-(x_1 - x_2)^2]. \quad (37)$$

Realizations of the field and/or the derivative of the field are measured at locations $x = 0$ and $x = 1$. The following four subcases of realizations are considered:

(i) Only the realizations of the field are measured at $x = 0$ and $x = 1$, i.e., $g(0) = \alpha$, $g(1) = \beta$. From Eq. (34) we have

$$\begin{bmatrix} 1 & e^{-1} \\ e^{-1} & 1 \end{bmatrix} \begin{Bmatrix} \lambda_1(x) \\ \lambda_2(x) \end{Bmatrix} = \begin{Bmatrix} e^{-x^2} \\ e^{-(x-1)^2} \end{Bmatrix} \quad (38)$$

which leads to

$$\lambda_1(x) = \frac{e^{-x^2} - e^{-1}e^{-(x-1)^2}}{1 - e^{-2}},$$

$$\lambda_2(x) = \frac{-e^{-1}e^{-x^2} + e^{-(x-1)^2}}{1 - e^{-2}}. \quad (39)$$

The conditional mean and variance of the field are obtained, respectively, as

$$E[G(x)|\, g(0) = \alpha;\, g(1) = \beta] = \frac{(\alpha - \beta e^{-1})e^{-x^2} + (\beta - \alpha e^{-1})e^{-(x-1)^2}}{1 - e^{-2}}$$

$$\text{Var}\,[G(x)|\, g(0) = \alpha;\, g(1) = \beta] = \sigma_0^2\left[1 - \frac{e^{-2x^2} - 2e^{-2(x^2-x+1)} + e^{-2(x-1)^2}}{1 - e^{-2}}\right]. \quad (40)$$

(ii) Only the realizations of the derivative of the field are measured at $x = 0$ and $x = 1$, i.e., $g'(0) = \gamma$, $g'(1) = \delta$. From Eq. (15) we have

$$\begin{bmatrix} 2 & -2e^{-1} \\ -2e^{-1} & 2 \end{bmatrix} \begin{Bmatrix} \mu_1(x) \\ \mu_2(x) \end{Bmatrix} = \begin{Bmatrix} 2xe^{-x^2} \\ 2(x-1)e^{-(x-1)^2} \end{Bmatrix}. \quad (41)$$

The solution of the above equation is

$$\mu_1(x) = \frac{xe^{-x^2} + (x-1)e^{-1}e^{-(x-1)^2}}{1 - e^{-2}},$$

$$\mu_2(x) = \frac{xe^{-1}e^{-x^2} + (x-1)e^{-(x-1)^2}}{1 - e^{-2}}. \quad (42)$$

The conditional mean and variance of the field are then

$$E[G(x)|\, g'(0) = \gamma;\, g'(1) = \delta] = \frac{x(\gamma + \delta e^{-1})e^{-x^2} + (x-1)(\delta + \gamma e^{-1})e^{-(x-1)^2}}{1 - e^{-2}}$$

$$\text{Var}\,[G(x)|\, g'(0) = \gamma;\, g'(1) = \delta] = \sigma_0^2\left[1 - \frac{2x^2e^{-2x^2} - 4x(x-1)e^{-2(x^2-x+1)} + 2(x-1)^2e^{-2(x-1)^2}}{1 - e^{-2}}\right]. \quad (43)$$

(iii) The realization of the field at $x = 0$ and the realization of the derivative of the field at $x = 1$ are measured, i.e., $g(0) = \alpha$, $g'(1) = \delta$. Equation (15) reads

$$\begin{bmatrix} 1 & -2e^{-1} \\ -2e^{-1} & 2 \end{bmatrix} \begin{Bmatrix} \lambda_1(x) \\ \mu_2(x) \end{Bmatrix} = \begin{Bmatrix} e^{-x^2} \\ 2(x-1)e^{-(x-1)^2} \end{Bmatrix}. \quad (44)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The solution is

$$\lambda_1(x) = \frac{e^{-x^2} + 2(x-1)e^{-1}e^{-(x-1)^2}}{1 - 2e^{-2}},$$

$$\mu_2(x) = \frac{e^{-1}e^{-x^2} + (x-1)e^{-(x-1)^2}}{1 - 2e^{-2}}. \quad (45)$$

The conditional mean and variance of the field are then

$$E[G(x)|g(0) = \alpha; g'(1) = \delta] = \frac{(\alpha + \delta e^{-1})e^{-x^2} + (x-1)(\delta + 2\alpha e^{-1})e^{-(x-1)^2}}{1 - 2e^{-2}}$$

$$\text{Var}[G(x)|g(0) = \alpha; g'(1) = \delta] = \sigma_0^2\left[1 - \frac{e^{-2x^2} + 4(x-1)e^{-2(x^2-x+1)} + 2(x-1)^2e^{-2(x-1)^2}}{1 - 2e^{-2}}\right]. \quad (46)$$

(iv) The realizations of both the field and the derivative of the field are measured at $x = 0$ and $x = 1$, $g(0) = \alpha$, $g(1) = \beta$, $g'(0) = \gamma$, and $g'(1) = \delta$. The governing equations for kriging weights now become

$$\begin{bmatrix} 1 & e^{-1} & 0 & -2e^{-1} \\ e^{-1} & 1 & 2e^{-1} & 0 \\ 0 & 2e^{-1} & 2 & -2e^{-1} \\ -2e^{-1} & 0 & -2e^{-1} & 2 \end{bmatrix} \begin{Bmatrix} \lambda_1(x) \\ \lambda_2(x) \\ \mu_1(x) \\ \mu_2(x) \end{Bmatrix} = \begin{Bmatrix} e^{-x^2} \\ e^{-(x-1)^2} \\ 2xe^{-x^2} \\ 2(x-1)e^{-(x-1)^2} \end{Bmatrix}. \quad (47)$$

The solution of the above equations is

$$\lambda_1(x) = \frac{(1 - 3e^{-2} + 4xe^{-2})e^{-x^2} + e^{-1}(-3 + e^{-2} + 2x - 2xe^{-2})e^{-(x-1)^2}}{1 - 6e^{-2} + e^{-4}}$$

$$\lambda_2(x) = \frac{e^{-1}(-1 - e^{-2} - 2x + 2xe^{-2})e^{-x^2} + (1 + e^{-2} - 4xe^{-2})e^{-(x-1)^2}}{1 - 6e^{-2} + e^{-4}}$$

$$\mu_1(x) = \frac{(2e^{-2} + x - 3xe^{-2})e^{-x^2} + e^{-1}(-2 + x + xe^{-2})e^{-(x-1)^2}}{1 - 6e^{-2} + e^{-4}}$$

$$\mu_2(x) = \frac{e^{-1}(1 - e^{-2} + x + xe^{-2})e^{-x^2} + (-1 + e^{-2} + x - 3xe^{-2})e^{-(x-1)^2}}{1 - 6e^{-2} + e^{-4}}. \quad (48)$$

The conditional mean and variance of the actual field are then

$$E[G(x)|g(0) = \alpha; g(1) = \beta; g'(0) = \gamma; g'(1) = \delta] = \alpha\lambda_1(x) + \beta\lambda_2(x) + \gamma\mu_1(x) + \delta\mu_2 \quad (x)$$

$$\text{Var}[G(x)|g(0) = \alpha; g(1) = \beta; g'(0) = \gamma; g'(1) = \delta] = \sigma_0^2\left[1 + \frac{f(x)}{1 - 6e^{-2} + e^{-4}}\right] \quad (49)$$

where

$$f(x) = (-3 + e^{-2} + 4x - 4xe^{-2} - 2x^2 + 6x^2e^{-2})e^{-2(x-1)^2} + (-1 + 3e^{-2} - 8xe^{-2} - 2x^2 + 6x^2e^{-2})e^{-2x^2} + (6 - 2e^{-2} + 4x + 4xe^{-2} - 4x^2 - 4x^2e^{-2})e^{-2(x^2-x+1)}. \quad (50)$$

Figure 1 portrays the conditional variances of the field for above four cases. Figure 2 gives amplified conditional variance for case 4 in the interval [0, 1]. It is seen that the conditional variance vanishes at those locations where the realizations of the field are measured; moreover, the slope of the conditional variance vanishes where the realizations of the derivative of the field are measured. As expected, both the conditional variance and its slope vanish where the realization of the field and its derivative are measured. A peak appears at the location where the realization of the derivative of the field vanishes. Far from the measured locations, the conditional variance tends to unconditional variance. This feature is analytically deduced from Eqs. (40), (43), (46), and (49). For cases 1, 2, and 4, the condi-

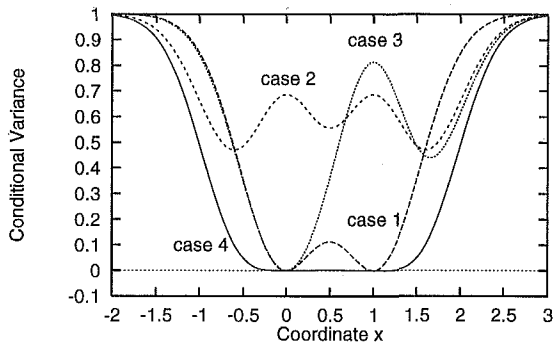


Fig. 1 The conditional variance of the field $G(x)$, normalized by the unconditional variance σ_0^2

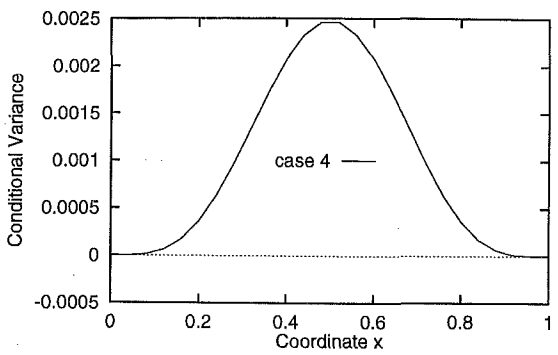


Fig. 2 Magnification of the conditional variance of the field $G(x)$ for case 4 in interval [0, 1], normalized by the unconditional variance σ_0^2

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

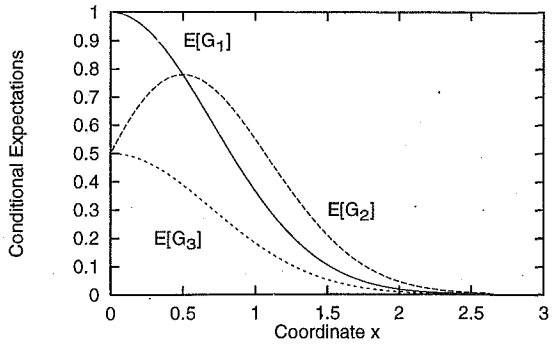


Fig. 3 Conditional means of the 3-variate field with realizations specified at $x = 0$

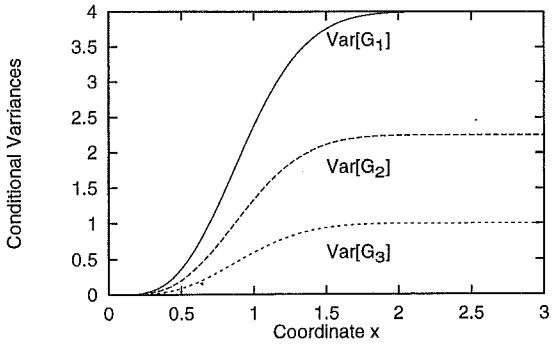


Fig. 4 Conditional variances of the 3-variate field with realizations specified at $x = 0$

tional variance is symmetric with respect to $x = 0.5$, since the realizations are symmetric and the field is homogeneous. It is seen that the realizations of the derivative of the field have significant effect on the conditional variance. The examination of Fig. 1 leads to a remarkable conclusion that conditioning of the derivative of the field significantly alters its properties of the field in the vicinity where the measurements have been made. Therefore, for sufficiently accurate conditional simulation, maximum information should be sought and utilized. It is suggested that realizations of the field as well as of its derivative should be measured wherever possible, in order to obtain more accurate estimation of the field at unmeasured locations.

Example 2

Consider a 3-component random field $\mathbf{G}(x) = [G_1(x), G_2(x), G_3(x)]^T$ with zero mean vector. The cross-covariance matrix of $\mathbf{G}(x)$ is assumed to be

$$C_{GG}(x_1, x_2) = \begin{bmatrix} \sigma_1^2 & r_{12}\sigma_1\sigma_2 & r_{13}\sigma_1\sigma_3 \\ r_{12}\sigma_1\sigma_2 & \sigma_2^2 & r_{23}\sigma_2\sigma_3 \\ r_{13}\sigma_1\sigma_3 & r_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix} \exp(-\xi^2) \quad (51)$$

where $\xi = x_2 - x_1$, σ_1, σ_2 and σ_3 are standard deviations, and r_{12}, r_{13}, and r_{23} are correlation coefficients. Assume that a set of realizations of $\mathbf{G}(x)$ and $\mathbf{G}'(x)$ at $x = 0$ are measured to be

$$g_1(0) = 1, \quad g_1'(0) = 0, \quad g_2(0) = 0.5$$

$$g_2'(0) = 1, \quad g_3(0) = 0.5, \quad g_3'(0) = 0. \quad (52)$$

With $n = 3$ and $N = M = 1$, Eq. (15) is comprised of six equations governing six kriging weights λ_{ml} and μ_{ml} ($m = 1, 2, 3$) for each component $G_l(x)$ of $\mathbf{G}(x)$. After solving out kriging weights λ_{ml} and μ_{ml}, the covariance matrix of the error vector is obtained from Eq. (22). The conditional means, variances, and covariances of the field are then obtained by using Eq. (27) and properties of the error vector. For the case $\sigma_1 = 2$, $\sigma_2 = 1.5$, $\sigma_3 = 1$, $r_{12} = 0.75$, $r_{13} = 0.25$, and $r_{23} = 0.5$, the conditional means, variances, and covariances of and between components of the field are illustrated in Figs. 3–5. It is seen that the conditional mean of each component starts from the realization value at $x = 0$ with the slope equal to the realization value of the derivative of the component. Comparison of the conditional means of $G_1(x)$ and $G_2(x)$, which have the same realization of the field but different realizations of the derivative, shows that the realizations of the derivatives of the field have significant effects on the conditional means. The conditional variances and covariances, however, are independent of the values of the realizations of the field, a conclusion which is identical to one obtained by Hoshiya (1994) and Kameda and Morikawa (1994). Here, however, we arrive at an additional feature: The conditional variances and covariances are also independent of the values of the realizations of the derivatives. The conditional variances and covariances are zero where the realizations of the field are given. As expected, the conditional means and conditional variances and covariances converge to their unconditional counterparts far from the locations the realization of the field or its derivative being measured.

In this connection, some natural questions arise: In the absence of measured derivatives, can one use numerical values of the field to calculate the derivatives? The reply is negative. Indeed, one can use numerical values derived from the measured values of the field, but it would not improve the result. Consider, for example, one-dimensional field $G(x)$. If we know two realizations $g(x_1)$ and $g(x_2)$, the realization of the derivative $G'(x_1)$ can be approximated as $g'(x_1) = [g(x_2) - g(x_1)]/(x_2 - x_1)$. However, the realizations $g(x_1)$ and $g(x_2)$ have already been taken into account to simulation the field. $G(x)$ conditioned by $g(x_1)$, $g(x_2)$ and $g'(x_1)$ and $G(x)$ conditioned by $g(x_1)$ and $g(x_2)$ coincide when x_2 moves to x_1. One practical way to monitor the derivatives, as we have mentioned in the Introduction of this study, is to use the strain/stress sensors as well as the displacement sensors.

As we have demonstrated above, the inclusion of information on derivatives may considerably alter the behavior of simulated results. In the example considered, the conditional variance is reduced when conditioned on the field and its derivatives. The arising pertinent question reads: Is this dependent on the problem considered? We must digress that the conditional covariance and variance are independent of the realizations of both the field and its derivatives. The conditional mean, however, is dependent on the realizations. The conditional variance is reduced when conditioned on the field and its derivatives. As far as the scale of the random field is concerned, it does not alter the property of the conditioned field.

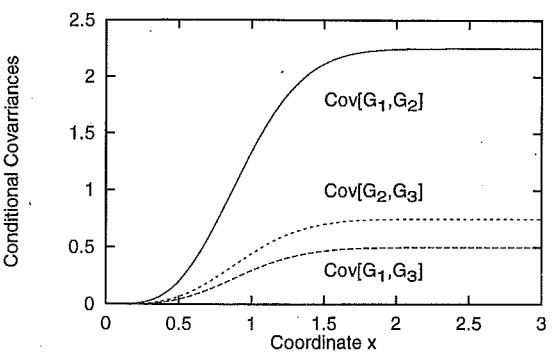


Fig. 5 Conditional covariances of the 3-variate field with realizations specified at $x = 0$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Conclusion

It is clearly demonstrated in this study that an information on the realizations of the derivatives of the random fields can considerably alter the behavior of simulated results. This implies that one should seek not only for the increased volume of statistical information on the fields themselves, but also for an additional information on their derivatives if possible. This additional information may prove to be extremely useful to make necessary decisions during the natural hazard monitoring for facilities with attendant high responsibility and/or public use.

Acknowledgment

This work has been supported by the National Center for Earthquake Engineering Research (NCEER), Buffalo, NY. Constructive criticisms of reviewers are gratefully appreciated.

References

Hoshiya, M., 1994, "Conditional Simulation of a Stochastic Field," *Structural Safety and Reliability,* Vol. 1, G. I. Schuëller, M. Shinozuka, and J. T. P. Yao, eds., Balkema, Rotterdam, pp. 349–353.

Hoshiya, M., and Maruyama, O., 1994, "Stochastic Interpolation of Earthquake Wave Propagation," *Structural Safety and Reliability,* Vol. 3, G. I. Schuëller, M. Shinozuka, and J. T. P. Yao, eds., Balkema, Rotterdam, pp. 2119–2124.

Ishii, K., and Suzuki, M., 1989, "Stochastic Finite Element Analysis for Spatial Variations of Soil Properties Using Kriging Technique," *Structural Safety and Reliability,* Vol. 1, A. H. S. Ang, M. Shinozuka, and G. I. Schuëller, eds., San Francisco, pp. 1161–1168.

Jeulin, D., and Renard, D., 1992, "Practical Limits of the Deconvolution of Images by Kriging," *Microse. Microanal Microstr.,* Vol. 3, pp. 333–361.

Journel, A. G., and Huijbregts, Ch. J., 1978, *Mining Geostatistics,* Academic Press, New York.

Kameda, H., and Morikawa, H., 1992, "Interpolating Stochastic Process for Simulation of Conditional Random Fields," *Probabilistic Engineering Mechanics,* Vol. 7, pp. 242–254.

Kameda, H., and Morikawa, H., 1994, "Conditioned Stochastic Processes for Conditional Random Fields," *Journal of Engineering Mechanics,* Vol. 120, pp. 855–875.

Krige, D. O., 1951, "A Statistical Approach to Some Basic Mine Valuation Problems," *J. Chem. Met'all. Min. Soc.,* Vol. 52, pp. 119–139.

Pourier, C., and Tinawi, R., 1991, "Finite Element Stress Tensor Field Interpolation and Manipulation Using 3D Dual Kriging," *Computer & Structures,* Vol. 40, pp. 211–225.

Trochu, F., 1993, "A Contouring Program Based on Dual Kriging Interpolation," *Engineering with Computers,* Vol. 9, pp. 160–177.

Vanmarcke, E. H., and Fenton, G. A., 1991, "Conditioned Simulation of Local Fields of Earthquake Ground Motion," *Structural Safety,* Vol. 10, pp. 247–274.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stepwise Loading of Half-Spaces in Elliptical Contact

J. Jäger

Ernst-Friedrich-Str. 13,
76227 Karlsruhe,
Germany

A new analytical solution for Hertzian surfaces in contact under stepwise oblique loading is presented in this paper. It is written as a superposition of so-called Cattaneo-Mindlin functions, which represent the tangential stress distribution for constant normal and monotonically increasing tangential forces. The size of the stick zone is determined by cones in the force space, which are also known as yield cones. It has a simpler form than the differential force-displacement relations of plasticity theory and the numerical discretization of the influence integrals. A short computer algorithm is proposed and compared with a numerical solution of the discretized problem.

1 Introduction

The great improvements of computer technology in the last years can replace highly difficult analytical investigations, but analytical formulas still have the advantage to show the influence of parameters directly and are easily evaluated, as long as they are simple enough. In this paper, an analytical model for the frictional contact of elastic bodies is proposed, which has been compared with a large number of numerical calculations. This model is very efficient for impact calculations, where the numerical solution is very computer-intensive (Jäger, 1994b, 1994c).

The theories of frictionless elastic contact by Hertz (1882), and tangential loading with friction by Cattaneo (1938) and Mindlin (1949), are outlined in Section two. They found that, in the case of a monotonically increasing tangential force with a constant normal force, the contact area is split into and inner, elliptical area of adhesion and an outer, annular region of partial slip. Due to the similarity between friction and plasticity problems, a differential formulation for two-dimensional loading was developed by Mindlin and Deresiewicz (1953). Three-dimensional loading was found important in a publication by Deresiewicz (1958), but the Mindlin-Deresiewicz model becomes so difficult that the first numerical solution was published only recently by Seridi and Dobry (1984) and Dobry and Seridi (1991). In that paper, a geometrical, differential relationship between conical yield surfaces is proposed, which determines the load history. A similar model was proposed by Iwan (1966) for a system of parallel elastic springs and Coulomb slip elements. Theoretical and numerical simulations for three-dimensional loading histories have also been reported in a paper by Chiang and Beck (1994), and similar contact laws have been found by Dundurs and Comninou (1981, 1983) and Comninou and Dundurs (1982) for the problem of an unbounded solid with a semi-infinite cut. The geometrical results of these studies are similar to the model for tangential loading presented in Section three, and torsional loading in Section four. A generalization for torsion of elliptical contact areas is reported in Jäger (1992).

In a much-quoted paper about tangential impact by Maw et al. (1975), the tangential pressure at radius r was approximated by special polynomials, which depend on the radius r and the semi-axis a of the contact area. Insertion of these polynomials in the half-space integrals (Jäger, 1992) gives the displacements in discrete form. The authors found that the method of Mindlin and Deresiewicz (1953) would require the knowledge of all traction increments of the previous load history and, therefore, they abandoned the Mindlin and Deresiewicz approach, because it places a limit on the number of increments. In fact, the complete load history is recorded in the program by Seridi and Dobry (1984). In contrast to the differential approach, we propose an incremental model in Section three, which consists of a superposition of Cattaneo-Mindlin solutions. Tangential and normal components of the loading increment are applied simultaneously, with monotonically increasing or decreasing values in constant direction. In earlier publications (Jäger, 1992, 1993, etc.) the loading was applied componentwise and the order of application was important. If the tangential forces are large enough, an elliptical slip area forms, and the old stress distribution is erased in the new slip area. Elimination of all erasing increments in the actual load history defines an equivalent load history, which produces the same stress distribution. During oblique loading, the three contact regimes of *complete overlapping, partial overlapping,* and *no overlapping* can be distinguished, depending on the size of the slip region of each increment. An analytical formula for the equivalent load history is obtained, which can easily be programmed, as illustrated in Section five. A numerical method for the solution of the load-displacement integrals of half-spaces by Jäger (1992) was used to test the analytical contact law for special examples in Section six. The assumption of constant slip directions in the slip area was not confirmed, but the global forces and displacements are reliable. For scientific purposes, all programs used in this paper are gratis available from the author.

2 Contact Laws by Hertz, Cattaneo, and Mindlin

If two bodies are pressed together by a normal force F_z, a normal pressure p_H acts on an elliptical contact area with the semi-axes a and b in x- and y-direction (Hertz, 1882)

$$p_H(a) = \begin{cases} \dfrac{3F_z}{2\pi ab}\sqrt{1-\rho^2}, & \text{for } [\rho \le 1], \\ 0, & \text{for } [\rho \ge 1], \end{cases}$$

$$\rho = \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}}, \quad (1)$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the International Mechanical Engineering Congress and Exposition, the Winter Annual Meeting of The ASME, Atlanta, GA, Nov. 17–22, 1996.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Sept. 12, 1994; final revision, Mar. 6, 1996. Associate Technical Editor: X. Markenscoff.

Paper No. 96-WA/APM-1.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where we introduced the dimensionless variable ρ. Three geometrical constants α, β, γ (Kalker, 1990), which depend on the ratio a/b or $A/(A+B)$, define the relation between the contact radius a, the normal force F_z, the semi-axes a/b and the curvatures A, B of the surface

$$\alpha = \frac{a^3(A+B)}{F_z(\vartheta_1+\vartheta_2)}, \quad \vartheta_i = \frac{(1-\nu_i)}{2G_i},$$
$$\beta = \frac{a^2(A+B)}{\zeta}, \quad \gamma = \frac{A}{A+B}, \tag{2}$$

where we introduced the shear module G_i, Poisson's number ν_i of body i and the combined displacement ζ of the bodies in the center of the contact area. A concise statement of all important formulas can be found in Jäger (1994a).

After application of the normal force, a tangential force F_x in x-direction and a balancing moment are mutually applied on the bodies. This problem was solved by Cattaneo (1938) for elliptical contact areas, and independently by Mindlin (1949) for circular contact areas. An inner area of adhesion forms, which is similar to the contact area, with the semi-axes a^* and b^*. We introduce the so-called Cattaneo-Mindlin functions $\sigma_{CM}(a, a^*)$, which are a superposition of Hertz functions $p_H(a)$ and $p_H(a^*)$ in tangential direction

$$\sigma_{CM}(a, a^*) = p_H(a) - p_H(a^*). \tag{3}$$

A normal pressure distribution of the amount of $p_H(a^*)$ corresponds to a contact area, which has the size of the stick area and the semi-axes a^*, b^*

$$p_H(a^*) = \begin{cases} \dfrac{3F_z^*}{2\pi a^* b^*}\sqrt{1-\rho^{*2}}, & \text{for} \quad \rho^* \le 1, \\ 0, & \text{for} \quad \rho^* \ge 1, \end{cases}$$
$$\rho^* = \sqrt{\frac{x^2}{a^{*2}} + \frac{y^2}{b^{*2}}}, \quad \frac{a^*}{b^*} = \frac{a}{b}. \tag{4}$$

The fictitious normal force F_z^* corresponds to a contact area of the size of the stick area. From the definition (1) for ρ and Eq. (4) follows $\rho/\rho^* = a^*/a$. Cattaneo (1938) assumed that a two-dimensional tangential stress distribution $\boldsymbol{\sigma}_t$, produced by a force $\mathbf{F}_t = (F_x, F_y)^T$, is of the form (3)

$$\boldsymbol{\sigma}_t = \begin{pmatrix} \sigma_{zx} \\ \sigma_{zy} \end{pmatrix} = \mathbf{p}\sigma_{CM}(a, a^*), \quad \mathbf{p} = \begin{pmatrix} p \\ q \end{pmatrix}, \quad |\mathbf{p}| = f, \tag{5a}$$

$$\mathbf{F}_t = \begin{pmatrix} F_x \\ F_y \end{pmatrix} = \mathbf{p}(F_z - F_z^*), \tag{5b}$$

$$F_z = \frac{2}{3}c_z\zeta^{3/2}, \quad F_z^* = \frac{2}{3}c_z\zeta^{*3/2}, \quad \zeta^* = a^{*2}\frac{(A+B)}{\beta}, \tag{5c}$$

with the stress direction vector $\mathbf{p}$. For constant values of F_z^*, i.e., constant size of the stick area, the forces are placed on a so-called yield cone in the three-dimensional force space, with its apex on the point F_z^* and a symmetry axes parallel to the F_z-axis. The corresponding displacements ξ and η of the initial contact point in x and y direction, respectively, can be calculated by integration of the load displacement equations (Jäger, 1992). Summarizing, we obtain a linear relation between the scaled displacement $\mathbf{e}$ in tangential direction and the normal component ζ, which can be interpreted as a yield cone in the displacement space

$$\mathbf{e} = \begin{pmatrix} c_x\xi/c_z \\ c_y\eta/c_z \end{pmatrix} = \mathbf{p}(\zeta - \zeta^*), \tag{6a}$$

$$\xi = \xi_1 + \xi_2, \quad \eta = \eta_1 + \eta_2, \tag{6b}$$

with the displacements ξ_i and η_i in x- and y-direction of the center of the contact area of body i; the fictitious compression ζ^* belongs to a contact ellipse with the semi-axis a^* and b^*, and the stiffness coefficients c_x, c_y, and c_z are defined by in the paper by Jäger (1994a). The assumption of constant stress directions in Eqs. (5) is only correct for $\nu_i = 0$, because otherwise a stress component lateral to the force arises in the slip area due to the coupling between the tangential stress components in x and y direction. This component is smaller than five percent of the absolute value of the stress for circular contact areas under monotonically increasing tangential forces (Johnson, 1985) and can be neglected in most cases. Furthermore, the angle between the force and the displacement varies between the extreme values for complete adhesion and gross slip for elliptical contact areas. This effect can also be neglected for large stick areas (Jäger, 1992).

3 General Contact Law

The Cattaneo-Mindlin contact law requires a constant normal force and uniaxial tangential displacement increments. In order to arrive at a more general contact law, we study a special load history with the displacement increment $\Delta\boldsymbol{\xi}_1 = (0, 0, \Delta\zeta_1)$ in normal direction and, subsequently, an oblique displacement increment $\Delta\boldsymbol{\xi}_2 = (\Delta\xi_2, \Delta\eta_2, \Delta\zeta_2)$. Physically, the components of $\Delta\boldsymbol{\xi}_2$ are increased monotonically from zero to the final value with constant proportion. Mathematically, a load history with infinitesimal increments $\mathbf{d}\boldsymbol{\xi}$ is applied in constant direction. To simplify matters, we can assume that the normal component $d\zeta$ is applied before the tangential increment $\mathbf{d}\boldsymbol{\xi}_t = (d\xi, d\eta)$. If the tangential increment is small enough, the new slip zone will always be smaller than the infinitesimal increase of the contact area, i.e., we have complete adhesion and obtain the tangential shear stress from (5a)

$$\mathbf{d}\boldsymbol{\sigma}_t = \mathbf{p}\sigma_{CM}(a, a^*)$$
$$= \mathbf{p}\frac{3f(A+B)a}{2\pi\alpha(\vartheta_1+\vartheta_2)b}(a\sqrt{1-\rho^2} - a^*\sqrt{1-\rho^{*2}}), \tag{7}$$

where the stress direction vector $\mathbf{p}$ is held constant and the factor before the brackets is a geometrical constant. For infinitesimal small slip areas, we perform the limit $a^* = a - da^*$ and obtain with the help of Eq. (4)

$$\mathbf{d}\boldsymbol{\sigma}_t = \mathbf{p}\frac{3f(A+B)a}{2\pi\alpha(\vartheta_1+\vartheta_2)b}\frac{da^*}{\sqrt{1-\rho^2}}. \tag{8}$$

For complete adhesion, we can set $da^* = da$, and obtain the value of the tangential stress by integration of (8). Comparison with (5a) shows, that the total increment of the tangential force at the end of step 2 can be written in form of a Cattaneo-Mindlin function

$$\Delta\boldsymbol{\sigma}_{t2} = \mathbf{p}_2\sigma_{CM}(a_2, a_1), \quad \text{for} \quad |\mathbf{p}_2| < f. \tag{9}$$

Formally, we can write $a_2^* = a_1$ in (5a). At the end of step 2 the entire contact area sticks completely and the stress distribution (9) is *frozen* in the stick area. This is the contact regime of *no overlapping* and no energy is lost by frictional forces. In the next step we superpose another monotonically increasing displacement increment $\Delta\boldsymbol{\xi}_3$ in constant direction, but now the tangential displacement may be so large that a slip area forms, i.e., $|\Delta\boldsymbol{\sigma}_{t3}| > f\,\Delta\sigma_{z3}$. If the new semi-axis of the stick area is denoted a_3^*, the solution of the boundary value problem is easily obtained by superposition of two Cattaneo-Mindlin functions: The first function $-\mathbf{p}_2\sigma_{CM}(a_2, a_3^*)$ makes the new slip area traction-free, and a second CM-function $\mathbf{p}_3\sigma_{CM}(a_3, a_3^*)$ satisfies Coulomb's law

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

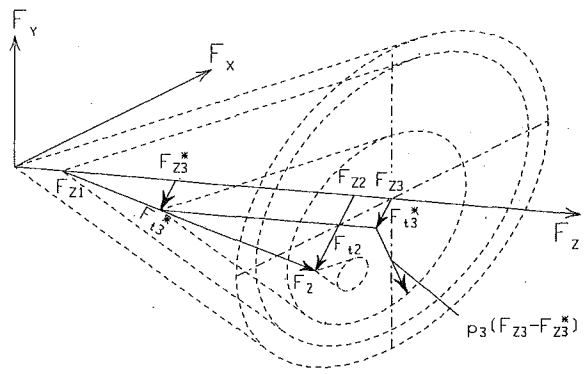


Fig. 1 Force cones with the slope $\tan^{-1} f$ from Eq. (11) in the force space define the contact regime

$$\Delta\boldsymbol{\sigma}_{t3} = \mathbf{p}_3\sigma_{CM}(a_3, a_3^*) - \mathbf{p}_2\sigma_{CM}(a_2, a_3^*),$$

$$|\mathbf{p}_3| = f, \quad \text{for} \quad a_2 \geq a_3^* \geq a_2^* = a_1. \quad (10)$$

We call the contact regime of Eq. (10) *partial overlapping,* because the new slip area overlays a part of the old contact area, i.e., $a_3^* \leq a_2$. We obtain the tangential force $\mathbf{F}_t$ by integration of Eqs. (9) and (10) similar to (5*b*)

$$\Delta\mathbf{F}_{t2} = \mathbf{p}_2(F_{z2} - F_{z1}), \quad (11a)$$

$$\Delta\mathbf{F}_{t3} = \mathbf{p}_3(F_{z3} - F_{z3}^*) - \mathbf{p}_2(F_{z2} - F_{z3}^*). \quad (11b)$$

The fictitous force F_{z3}^* corresponds to a contact area with the semiaxes a_3^*, b_3^* of the actual stick area. To every CM-function belongs a scaled displacement $\mathbf{e}$ from (6*a*)

$$\Delta\mathbf{e}_2 = \mathbf{p}_2(\zeta_2 - \zeta_1), \quad (12a)$$

$$\Delta\mathbf{e}_3 = \mathbf{p}_3(\zeta_3 - \zeta_3^*) - \mathbf{p}_2(\zeta_2 - \zeta_3^*). \quad (12b)$$

Figure 1 illustrates the relation (11*a*) and (11*b*) between the force components in the force space. All force cones have the slope $\tan^{-1} f$ and a symmetry axes parallel to the *z*-axes. Force cones with their apex at a force point $\mathbf{F}_i$ enclose regions of new increments $\Delta\mathbf{F}_{i+1}$ with complete adhesion. If the force is moved on a force cone with its apex at $\mathbf{F}_3^*$, the size of the stick area is constant and only the stress direction $\mathbf{p}_i$ in the slip area varies perpendicular to the cone, like in Fig. 1. All increments outside of the force cone with its apex at the origin generate complete sliding. The first increment $\Delta\boldsymbol{\xi}_1 = (0, 0, \Delta\zeta_1)$ is applied in normal direction, so that the force point arrives at $\mathbf{F}_1 = (0, 0, F_{z1})$. In the next step, the force is increased along the straight line from $\mathbf{F}_1$ to $\mathbf{F}_2$ inside of the force cone at $\mathbf{F}_1$, i.e., complete adhesion takes place. The new force increment $\Delta\mathbf{F}_3$ is located outside of the next cone with its apex at $\mathbf{F}_2$, partial overlapping takes place and the new force point is placed on a force cone, which has its apex at $\mathbf{F}_3^*$ on the path $\mathbf{F}_1 - \mathbf{F}_2$, as shown in Fig. 1. All forces on the force cone $\mathbf{F}_3^*$ have the same stick area with the semi-axes a_3^* and b_3^* and the old load history between $\mathbf{F}_2$ to $\mathbf{F}_3^*$ is completely erased by the new slip zone. An equivalent load history with the same stress distribution can be found, if the second increment is applied from $\mathbf{F}_1$ to $\mathbf{F}_3^*$, followed by an oblique increment on the force cone from $\mathbf{F}_3^*$ to $\mathbf{F}_3$. If the force $\mathbf{F}_3^*$ reduces to the value $\mathbf{F}_1$, the "frozen" stress distribution $\Delta\boldsymbol{\sigma}_{t2}$ is completely erased and the stress distribution would be the same after application of $\Delta\mathbf{F}_2 + \Delta\mathbf{F}_3$ in one step. If a new force increment $\Delta\mathbf{F}_4$ is applied, the procedure can be repeated with the equivalent load history, i.e., complete adhesion takes place if the increment is enclosed by the force cone with its apex at $\mathbf{F}_3$, etc. The displacement cones in the displacement space are identical with the force cones in Fig. 1, after replacing the force components (F_{xi}, F_{yi}, F_{zi}) by the displacements (ξ_i, η_i, ζ_i).

The general solution can be written as a superposition of CM-functions

$$\boldsymbol{\sigma}_{zt} = \sum_{i=1,2,3\ldots}^{k} \Delta\boldsymbol{\sigma}_i, \quad (13a)$$

$$\boldsymbol{\sigma}_{zt} = \begin{pmatrix} \sigma_x \\ \sigma_y \end{pmatrix}, \quad \Delta\boldsymbol{\sigma}_i = \begin{pmatrix} \Delta\sigma_{xi} \\ \Delta\sigma_{yi} \end{pmatrix}, \quad \mathbf{p}_i = \begin{pmatrix} p_i \\ q_i \end{pmatrix}, \quad (13b)$$

$$\Delta\boldsymbol{\sigma}_i = \begin{cases} \mathbf{p}_i\sigma_{CM}(a_i, \ a_i^*), \quad a_i^* = a_{i-1}, |\mathbf{p}_i| < f, \\ \qquad\qquad\qquad\qquad\qquad \text{(no overlapping)}, \quad (13c) \\ \mathbf{p}_i\sigma_{CM}(a_i, a_i^*) - \mathbf{p}_{i-1}\sigma_{CM}(a_{i-1}, a_i^*), |\mathbf{p}_i| = f, \\ \quad a_{i-1} \geq a_i^* \geq a_{i-1}^*, \quad \text{(partial overlapping)}. \quad (13d) \end{cases}$$

The upper summation index k and the stress increment $\Delta\boldsymbol{\sigma}_i$ of Eq. (13*a*) depend on the radius a_i^* of the new stick area. For complete overlapping, the old stress increment is erased and the increment at step $i-1$ must be replaced by the sum of increments of step i and step $i-1$

$$(\Delta\xi, \Delta\eta, \Delta\zeta)|_{\text{step}i-1} \leftarrow (\Delta\xi, \Delta\eta, \Delta\zeta)|_{\text{step}i-1} + (\Delta\xi, \Delta\eta, \Delta\zeta)|_{\text{step}i},$$

$$\text{for } a_i^* \leq a_{i-1}^*, (\Delta\xi, \Delta\eta, \Delta\zeta)|_{\text{step}i} \leftarrow (0, 0, 0). \quad (14)$$

This procedure eliminates all increments of complete overlapping, so that only partial overlapping and no overlapping remain. In the case of no overlapping in (13*c*), the stress increments are linearly superposed. Otherwise, that part where the old and the new contact area overlap must be made traction-free by the function $-\mathbf{p}_{i-1}\sigma_{CM}(a_{i-1}, a_i^*)$, and superposition of another Cattaneo-Mindlin function $\mathbf{p}_i\sigma_{CM}(a_i, a_i^*)$ satisfies Coulomb's law in the slip area and performs a rigid-body displacement of the stick area. From Eq. (5*b*) follows the force

$$\mathbf{F}_t = \begin{pmatrix} F_x \\ F_y \end{pmatrix} = \sum_{i=1,2,3\ldots}^{k} \Delta\mathbf{F}_{ti}, \quad \Delta\mathbf{F}_{ti} = \begin{pmatrix} \Delta F_{xi} \\ \Delta F_{yi} \end{pmatrix}, \quad (15a)$$

$$\Delta\mathbf{F}_{ti} = \begin{cases} \mathbf{p}_i(F_{z,i} - F_{z,i-1}), \quad F_{z,i}^* = F_{z,i-1}, \\ \qquad\qquad\qquad\qquad \text{for} \quad |\mathbf{p}_i| < f, \quad (15b) \\ \mathbf{p}_i(F_{z,i} - F_{z,i}^*) - \mathbf{p}_{i-1}(F_{z,i-1} - F_{z,i}^*), |\mathbf{p}_i| = f, \\ \qquad\qquad \text{for} \quad \zeta_{i-1} \geq \zeta_i^* \geq \zeta_{i-1}^*. \quad (15c) \end{cases}$$

A similar equation follows from (6*a*) and (13) for the displacements

$$\mathbf{e}_t = \begin{pmatrix} c_x\xi/c_z \\ c_y\eta/c_z \end{pmatrix} = \sum_{i=1,2,3\ldots}^{k} \Delta\mathbf{e}_i, \quad \Delta\mathbf{e}_i = \begin{pmatrix} c_x\Delta\xi_i/c_z \\ c_y\Delta\eta_i/c_z \end{pmatrix}, \quad (16a)$$

$$\Delta\mathbf{e}_i = \begin{cases} \mathbf{p}_i(\zeta_i - \zeta_{i-1}), \quad \text{for} \quad |\mathbf{p}_i| < f, \quad (16b) \\ \mathbf{p}_i(\zeta_i - \zeta_i^*) - \mathbf{p}_{i-1}(\zeta_{i-1} - \zeta_i^*), \quad |\mathbf{p}_i| = f, \\ \qquad\qquad \text{for} \quad \zeta_{i-1} \geq \zeta_i^* \geq \zeta_{i-1}^*. \quad (16c) \end{cases}$$

For a given increment $\Delta\boldsymbol{\xi}_i$ or $\Delta\mathbf{F}_i$, respectively, the value of ζ_i^* and F_{zi}^* can be determined from (15) and (16)

$$\mathbf{p}_i = \frac{\Delta\mathbf{e}_i}{\zeta_i - \zeta_{i-1}} = \frac{\Delta\mathbf{F}_{t,i}}{F_{z,i} - F_{z,i-1}}, \quad \text{for} \quad |\mathbf{p}_i| < f, \quad (17a)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\zeta_i^* = \zeta_i - \left\{ \sqrt{\left(\frac{\Delta \mathbf{e}_i' \mathbf{p}_{i-1}}{\Delta e_i'^2} \right)^2 + \frac{f^2 - \mathbf{p}_{i-1}^2}{\Delta \mathbf{e}_i'^2}} - \frac{\Delta \mathbf{e}_i' \mathbf{p}_{i-1}}{\Delta \mathbf{e}_i'^2} \right\}^{-1}, \tag{17b}$$

$$F_{z,i}^* = F_{z,i} - \left\{ \sqrt{\left(\frac{\Delta \mathbf{F}_{ti}' \mathbf{p}_{i-1}}{\Delta \mathbf{F}_{ti}'^2} \right)^2 + \frac{f^2 - \mathbf{p}_{i-1}^2}{\Delta F_{ti}'^2}} - \frac{\Delta \mathbf{F}_{ti}' \mathbf{p}_{i-1}}{\Delta \mathbf{F}_{ti}'^2} \right\}^{-1}, \quad \zeta_i^* \geq \zeta_{i-1}^*. \tag{17c}$$

$$|\mathbf{p}_i| = f, \quad \Delta \mathbf{e}_i' = \Delta \mathbf{e}_i - \mathbf{p}^{i-1} \Delta \zeta_i, \quad \Delta \mathbf{F}_{ti}' = \Delta \mathbf{F}_{ti} - \mathbf{p}_{i-1} \Delta F_{zi}, \tag{17d}$$

In the case of $|\mathbf{p}_{i-1}| = f$, we have $\Delta \mathbf{e}_i' \mathbf{p}_{i-1} < 0$ and $\Delta \mathbf{F}_{ti}' \mathbf{p}_{i-1} < 0$, so that the expression in brackets of (17) is always different from zero. For $\Delta \mathbf{e}_i' = 0$, elastic unloading occurs with $\zeta_i^* = \zeta_i$, i.e., during unloading no dissipation takes place and the unloading path follows exactly the loading path.

The differential compliances (Mindlin, 1953) or flow rules (Jarzębowskij et al., 1994) are more complicated than the expressions (15) and (16). As example, two equal spheres are first compressed by the normal force ΔF_{z1}. Than, the increment $\Delta \mathbf{F}_2 = (\Delta F_{x2}, 0, \Delta F_{z2})^T$, $\Delta F_{x2} > 0$ is applied, so that partial overlapping of the old contact area $\rho \leq a_1$ takes place. Next, an increment $\Delta \mathbf{F}_3 = (\Delta F_{x3}, 0, \Delta F_{z3})^T$, $\Delta F_{x3} < 0$, generates reverse slip with partial overlapping. The force component in y-direction is always zero: $\Delta F_{yi} = 0$. We obtain the formula for the tangential displacement of step 2 from (16c) with (2), (5c) and $\mathbf{p}_1 = 0$, $|\mathbf{p}_2| = f$

$$\frac{c_x}{c_z} \xi_2 = f \left(\frac{3}{2c_z} \right) \left\{ F_{z2}^{2/3} - \left(F_{z2} - \frac{F_{x2}}{f} \right)^{2/3} \right\}. \tag{18}$$

The geometrical parameters (2) are (Jäger, 1994a)

$$A + B = \frac{2}{R_1}, \quad \alpha = \frac{3}{4}, \quad \beta = 1,$$

$$\vartheta_1 + \vartheta_2 = \frac{1 - \nu_1}{G_1}, \quad \delta = \frac{1}{2},$$

$$\frac{3}{2c_z} = \frac{3(1 - \nu_1)}{2G_1 \sqrt{2R_1}}, \quad \frac{c_x}{c_z} = \frac{2(1 - \nu_1)}{2 - \nu_1}. \tag{19}$$

Differentiation of (18) and insertion of (19) yields

$$\frac{d\xi_2}{dt} = \frac{f}{c_x} \left(\frac{1}{\sqrt{\zeta_2}} - \frac{1}{\sqrt{\zeta_2^*}} \right) \frac{dF_{z2}}{dt} + \frac{1}{c_x \sqrt{\zeta_2^*}} \frac{dF_{x2}}{dt}. \tag{20}$$

Substitution of F_{zi} for ζ_i in (20) gives

$$\frac{d\xi_2}{dt} = \frac{2 - \nu_1}{4G_1} \frac{1}{a_2} \left\{ f \left[1 - \left(1 - \frac{F_{x2}}{f F_{z2}} \right)^{-1/3} \right] \frac{dF_{z2}}{dt} + \left(1 - \frac{F_{x2}}{f F_{z2}} \right)^{-1/3} \frac{dF_{x2}}{dt} \right\}. \tag{21}$$

Formula (21) is identical with the flow rule in the paper from Jarzębowski et al. (1994) for expanding slip zone. Similarly, we obtain the compliance for the third increment $\mathbf{p}_3 = (-f, 0)^T$

$$\frac{c_x}{c_z} \Delta \xi_3 = 2f \zeta_3^* - f(\zeta_3 + \zeta_2),$$

$$F_{z3}^* = \frac{2}{3} c_z \zeta^{*3/2} = \frac{\Delta F_{x3}}{2f} + \frac{F_{z3} + F_{z2}}{2}. \tag{22}$$

Partial differentiation of $\Delta \xi_3$ with respect to F_{x3} and F_{z3} gives the flow rule

$$\frac{d\xi_3}{dt} = \frac{f}{c_x} \left\{ \frac{1}{\sqrt{\zeta_3^*}} - \frac{1}{\sqrt{\zeta_3}} \right\} \frac{dF_{z3}}{dt} + \frac{1}{c_x \sqrt{\zeta_3^*}} \frac{dF_{x3}}{dt}. \tag{23}$$

For the special case $F_{z3} = F_{z2} =$ constant, expression (22) and (23) are identical with the formula by Jarzębowski (1994) for reverse slip. Consequently, the calculation of the compliances is possible, but the expressions depend on the load history and the formulas are much more difficult than the integral formulations (15) and (16).

4 Torsion of Axisymmetric Bodies

Application of a torsional couple M_z around the contact normal, i.e., the normal in the center of the common contact area, under a constant normal force, produces a tangential stress distribution with a circumferential component only (Lubkin, 1951). For axisymmetric bodies, an annular sliding area forms at the border of the contact area $a^* \leq \rho \leq a$, where Coulomb's law for full sliding must be satisfied, while the stick area must perform a rigid-body rotation. The solution was published by Lubkin (1951) and a generalization for arbitrary axisymmetric surfaces was reported by Jäger (1995). In the case of spheres, the torsional moment $M(a, a^*)$ and the combined twisting angle $\beta(a, a^*)$ can be approximated by

$$\frac{1}{f F_z a} M(a, a^*) = k^2 \left(1 - k^2 + \frac{3\pi}{16} k^2 \right),$$

$$k^2 = 1 - \frac{a^{*2}}{a^2}, \tag{24a}$$

$$\frac{a^2}{f F_z \left(\frac{1}{G_1} + \frac{1}{G_2} \right)} \beta(a, a^*) = \frac{-9\pi}{16^2} \ln \left[(1 - k^2) \left(1 + k^2 - \frac{16}{3\pi} k^2 \right) \right], \tag{24b}$$

with a maximum error of less than four percent (Jäger, 1993). Equations (24a) and (24b) hold for equal materials, monotonically increasing moments M and constant normal pressure. Numerical calculations in Jäger (1992) show, that these formulas

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

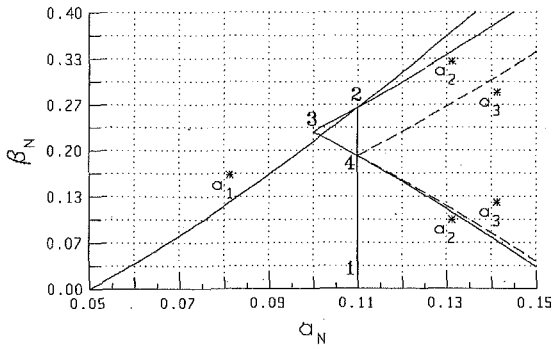


Fig. 2 Lines of constant size of the stick area for step i are denoted by a_i^*

provide a good approximation for elliptical contact areas and dissimilar materials. Substitution of the nondimensional constants M_N, β_N, and a_N

$$a_N = a\left(\frac{1}{R_1} + \frac{1}{R_2}\right), \quad M_N = M\frac{16a^3}{3\pi f F_z}\left(\frac{1}{R_1} + \frac{1}{R_2}\right)^4,$$

$$\beta_N = \beta\frac{16^2 a^3\left(\frac{1}{R_1} + \frac{1}{R_2}\right)}{9\pi f F_z\left(\frac{1}{G_1} + \frac{1}{G_2}\right)}, \tag{25}$$

in Eqs. (24) gives

$$M_N(a_N, a_N^*) = a_N^4 + a_N^2 a_N^{*2}\left(\frac{16}{3\pi} - 2\right) + a_N^{*4}\left(1 - \frac{16}{3\pi}\right), \tag{26a}$$

$$\beta_N(a_N, a_N^*) = -a_N \ln\left[\frac{a_N^{*2}}{a_N^2}\left\{2 - \frac{16}{3\pi} + \frac{a_N^{*2}}{a_N^2}\left(\frac{16}{3\pi} - 1\right)\right\}\right]. \tag{26b}$$

An example for Eqs. (26) is illustrated in Figs. 2 and 3 for values of $a_N \leq 0.2$. Lines marked with a_i^* define yield cones (26) with constant stick area, similar to the force cones in Fig. 1, but the generators are nonlinear. First, the bodies are compressed to the value a_1, marked as point 1 in the Figs. 2 and 3. Than, the torsional angle is monotonically increased, so that a stick radius a_1^* forms, marked as point 2 in the figures. Subsequently, the moment is reduced and a new sliding area with radius a_2^* overlays the old contact area, marked as point 4 in the figures. The moment at point 4 follows by subtraction of $M_N(a_1, a_2^*)$, which makes the new slip area $a_2^* \leq \rho \leq a_1$ traction-free, and subtraction of another moment $M_N(a_1, a_2^*)$, which satisfies Coulomb's law in the slip area and produces a constant rotation of the stick area. This boundary value problem can also be satisfied by a loading history following the line 2-3 for constant stick area a_2^*, which unloads the area $a_2^* \leq \rho \leq a_1$ from tangential traction and reduces the contact area to the size of a_2^*. The line 3-4 loads the sliding area by a negative tangential traction and increases the contact area to a_1. Point 2 is defined by $a_N = a_1$ and $a_N^* = a_1^*$ in (26) and point 3 by subtraction of $M_N(a_1, a_2^*)$ or $\beta_N(a_1, a_2^*)$. The lines a_3^* are the load paths for a new sliding area with constant size outside of the old sliding area. In contrast to Fig. 1, the yield cones are overlapping and not placed side by side. The process of superposition is identical with the model in Section three (Jäger, 1993),

$$M = \sum_{i=1}^{k} \Delta M_i, \tag{27a}$$

$$\Delta M_i = \begin{cases} M_N(a_i, a_i^*), & \text{for} \quad a_i^* \geq a_i \\ M_N(a_i, a_i^*) - M_N(a_{i-1}, a_i^*), & \\ & \text{for} \quad a_{i-1} \geq a_i^* \geq a_{i-1}^*, \end{cases} \tag{27b, 27c}$$

$$\beta = \sum_{i=1}^{k} \Delta\beta_i, \tag{28a}$$

$$\Delta\beta_i = \begin{cases} \beta_N(a_i, a_i^*), & \text{for} \quad a_i^* \geq a_i \\ \beta_N(a_i, a_i^*) - \beta_N(a_{i-1}, a_i^*), & \\ & \text{for} \quad a_{i-1} \geq a_i^* \geq a_{i-1}^*. \end{cases} \tag{28b, 28c}$$

Equations (27b) and (28b) define the contact regime of no overlapping, (27c) and (28c) partial overlapping, as described in Section 3. In the case of complete overlapping, the old and the new increment must be added and the increment with index i can be made zero

$$\Delta M_{i-1} \leftarrow \Delta M_{i-1} + \Delta M_i, \quad \Delta\beta_{i-1} \leftarrow \Delta\beta_{i-1} + \Delta\beta_i,$$
$$\Delta M_i \leftarrow 0, \quad \Delta\beta_i \leftarrow 0, \tag{29}$$

and process (27) must be repeated for the new values ΔM_{i-1} and $\Delta\beta_{i-1}$.

5 Computation of the Contact Law

The input values of the program are the consecutive force or displacement increments of a load history. After calculating the stress direction $\mathbf{p_i}$ from (17a), the inequality $|\mathbf{p_i}| < f$ must be checked. If it holds, we have the contact regime of no overlapping. Otherwise, we must evaluate the value of ζ_i^* from (17b). If the inequality $(\min(\zeta_i, \zeta_{i-1}) \geq \zeta_i^* \geq \zeta_{i-1}^*)$ of (17b) holds, we have partial overlapping; otherwise, complete overlapping takes place. In the last case, the increments of the last two steps must be added in the manner of Eq. (14), and the contact regime must again be determined. The output of the program is an equivalent load history, which produces the same stress distribution.

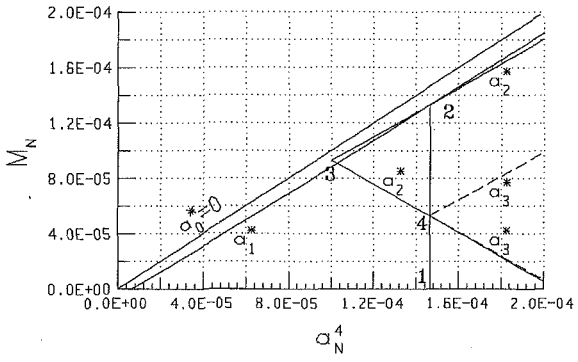


Fig. 3 The moment M_N for constant size of the stick area, marked with a_i^*

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

1 Choose input variables:
$f, G_1, G_2, \nu_1, \nu_2, a/b, A + B;$
2 Initialize
c_x, c_y, c_z from (Jäger, 1994a)
$\zeta_0 := \zeta_0^* := 0;\ \mathbf{p_0} = \Delta\boldsymbol{\xi}_0 = \Delta\mathbf{F_0} = \mathbf{0};$ Inc := 0;
3 Recursion is now carried out for all values of the load history
Repeat { Begin the outer Repeat-Until-Loop }
Inc := Inc + 1;
i := Inc;
Choose next increment: $\Delta\boldsymbol{\xi}_\mathbf{i}$ or $\Delta\mathbf{F}_\mathbf{i}$ for $1 \le i \le N$;
Repeat { Determine the contact regime }
Done := True;
Calculate $\mathbf{p}_\mathbf{i}$ from (17a)

$$\zeta_i = \begin{cases} \zeta_{i-1} + \Delta\zeta_i, & \{\text{displacement-controlled}\} \\ \dfrac{3}{2c_z}\left(F_{z,i-1} + \Delta F_{zi}\right)^{2/3} & \{\text{force-controlled}\} \end{cases}$$

If ($|\mathbf{p}_\mathbf{i}| \le f$) and ($\zeta_i \ge \zeta_{i-1}$) Then Begin { no overlapping }
$\zeta_i^* = \zeta_{i-1};$
End Else Begin
Calculate ζ_i^* from (17b) or (17c)
If ($\min(\zeta_i, \zeta_{i-1}) \ge \zeta_i^* \ge \zeta_{i-1}^*$) Then Begin { partial overlapping }
Determine $\mathbf{p}_\mathbf{i}$ from (15c) or (16c);
End Else Begin
If ($i > 1$) Then Begin { complete overlapping }
Done := False;
$\Delta\boldsymbol{\xi}_{\mathbf{i-1}} = \Delta\boldsymbol{\xi}_{\mathbf{i-1}} + \Delta\boldsymbol{\xi}_\mathbf{i};\ \Delta\mathbf{F}_{\mathbf{i-1}} = \Delta\mathbf{F}_{\mathbf{i-1}} + \Delta\mathbf{F}_\mathbf{i};$
$\Delta\boldsymbol{\xi}_\mathbf{i} := 0;\ \Delta\mathbf{F}_\mathbf{i} := 0;$
$i := i - 1;$

$$\zeta_i^* = \begin{cases} \zeta_{i-1} + \Delta\zeta_i, & \{\text{displacement-controlled}\} \\ \left(\zeta_{i-1}^{3/2} + \Delta F_{zi}\dfrac{3}{2c_z}\right)^{2/3}, & \{\text{force-controlled}\} \end{cases}$$

End Else Begin { full sliding }
$\zeta_1^* = 0;$
Determine $\Delta\boldsymbol{\xi}_\mathbf{1}$ and $\Delta\mathbf{F}_\mathbf{1}$ for full sliding from (5), (6);
End;
End;
End;
Until (Done = True); { Contact regime found }
For i := 1 to Inc do
Determine $\Delta\boldsymbol{\xi}_\mathbf{i}$ or $\Delta\mathbf{F}_\mathbf{i}$ resp., from (15) and (16);
Until (Inc $\ge N$); { Termination of the outer Repeat-Until-Loop }
4 Termination of the program.

The characteristic values of ζ_i, ζ_i^* and $\mathbf{p}_\mathbf{i}$ for the equivalent load history are now determined. For $\zeta_i^* \ge \zeta_i$ we have the contact regime of no overlapping and otherwise partial overlapping. The stress distribution can be determined from (13f), (13g), the force from (15b), (15c), and the displacement from (16b), (16c). In the case of complete overlapping, the actual increments $\Delta\boldsymbol{\xi}_\mathbf{i}$ and $\Delta\mathbf{F}_\mathbf{i}$ have been set zero in this program. In earlier publications, all increments have been shifted ($\Delta\boldsymbol{\xi}_\mathbf{i} \leftarrow \Delta\boldsymbol{\xi}_{\mathbf{i+1}}$), so that all zero entries disappear. The equivalent load history differs from the real load history, when increments of complete overlapping erased a part of the history, i.e., $\Delta\boldsymbol{\xi}_\mathbf{i} = \mathbf{0}$ for at least one increment.

6 Comparison With Numerical Results

The correct solution of the load displacement equation for half-spaces can only be solved numerically, with large nonlinear equation systems (e.g., matrices with $9*10^6$ elements for 1000 points), using iteration methods like the Gauss-Seidel method (Jäger, 1992). This method converges savely for displacement-controlled load histories. In Fig. 4 a load history for two equal spheres with Poisson's number $\nu_i = 0$ under constant normal compression is illustrated, where the displacement moves clockwise to the corners of the sectors of a circle. The sectors have a central angle of 22.5 deg and the numbers denote the corresponding load step. In Fig. 5, the tangential displacement is first increased to the value η_1, than decreased to the value η_2, and moved clockwise on a displacement or yield cone. Figures 6 and 7 show the corresponding force in the force space, which should also be placed on the corners of the sectors. The error of the present model is in most cases smaller than the error of the underlying Coulomb model, because the coefficient of friction depends on many parameters which are difficult to control, as the normal force, relative velocities, and humidity. Therefore,

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

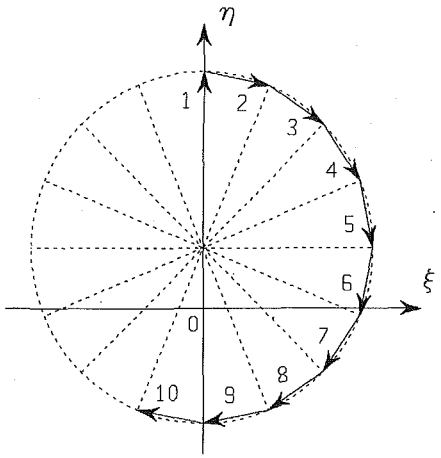


Fig. 4 Variation of ξ, η on a displacement cone for $\zeta = \zeta_1$

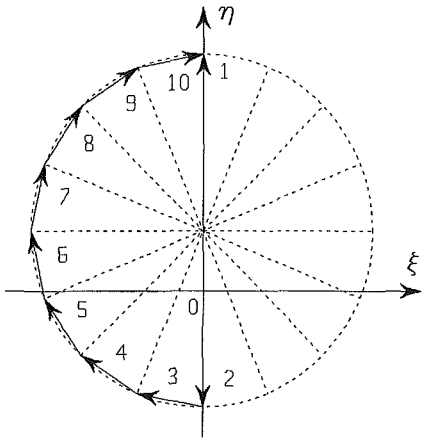


Fig. 5 Variation of ξ, η on a displacement cone after reverse slip, for $\zeta = \zeta_1$

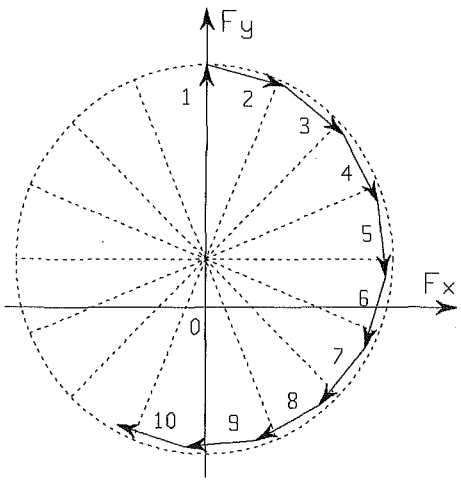


Fig. 6 The resulting force of the load history of Fig. 4

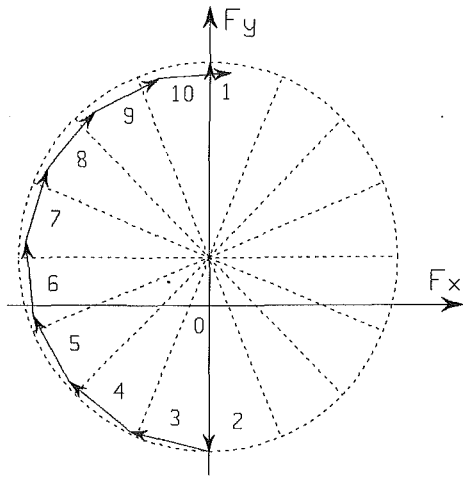


Fig. 7 The force corresponding to the load history of Fig. 5

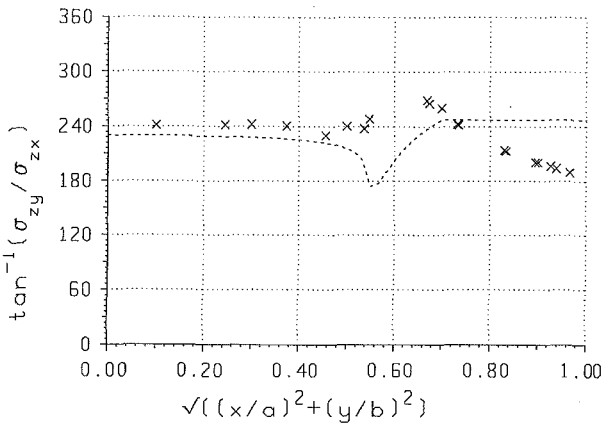


Fig. 8 The stress angle for step 10 of the load history of Fig. 4 is marked with crosses, compared with the analytical approximation (dashed)

a simple approach can be useful to interprete the complex numerical results. Coulomb described his approach as "a mixture of mathematics and physics" (Tabor, 1981). Its success is due to the simplicity, and the mathematical error found in experiments can often be neglected.

In Fig. 8 the stress angles are shown as a function of the nondimensional radius ρ. The analytical approximation, drawn with a broken line, is constant in the slip area $\rho \geq 0.7$. Comparison of both load histories shows, that the numerical stress angle (markers) depends strongly on the load history. It can be concluded that after application of a new increment in a different direction the stress reorientation starts at the border of the contact area and moves inwards with increasing values of the increment. The computer program of this article is available from the author as public domain program.

7 Conclusion

A new analytical formula for oblique loading of elastic bodies in contact has been proposed, which is easy to program. It is based on the assumption that the slip area slides like a rigid body and, both, the coupling of the tangential stress components in x- and y-direction for $\nu_i \neq 0$ as well as the coupling of normal and tangential stress components for dissimilar materials are neglected. It consists of a finite superposition of so-called Cattaneo-Mindlin functions, which is mathematically simpler than a differential formulation. It was illustrated geometrically by force and displacement cones, which enclose the regions of complete adhesion. The displacement cones are a special characteristic of spherical bodies, while the force cones are more general. After the description of the computer program, the contact law was compared with a numerical solution of the half-space integrals and a good agreement was found for the values of the forces and the stick area. The direction of the slip stress in the slip area is only correct for nearly uniaxial load

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

histories, which was not the case for the examples in this paper. This contact law is very efficient for frictional impact, fracture mechanics, vibration problems and contact of granular materials, where large computation times are necessary for numerical contact solutions.

Acknowledgment

The author is grateful to Professor Mróz and Doctor Jarzębowskij for inspirations and the discussion of the examples, during a short visit in Warshaw at the Institute of Fundamental Technology. He is also grateful to the DFG (German Research Community) for the evaluation of this project, although it was not supported.

References

Abramowitz, M., and Stegun, I., 1972, *Handbook of mathematical functions,* Dover, New York.

Cattaneo, C., 1938, "Sul contatto di due corpi elastici: distribuzione locale degli sforzi," *Rendiconti delle Academia nazionale dei Lincei,* Vol. 27, Ser. 6, 342–348; 434–436; 474–478.

Chiang, D. Y., and Beck, J. L., 1994, "A New Class of Distributed-Element Models for Cyclic Plasticity," *International Journal of Solids and Structures,* Vol. 31, No. 4, Part I: pp. 469–484; Part II: pp. 485–496.

Comninou, M., and Dundurs, J., 1982, "An Educational Elasticity Problem With Friction Part 2: Unloading for Strong Friction and Reloading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 47–51.

Deresiewicz, H., 1958, "Mechanics of Granular Matter," *Advances in Applied Mechanics,* Vol. 5, Dryden, Karman, Kuerti, eds., Academic Press, New York, pp. 233–306.

Dobry, R., Nt, T.-T., Petrakis, E., and Seridi, A., 1991, "General Model for Contact Law Between Two Rough Spheres," ASCE *Journal of Engineering Mechanics,* Vol. 117, pp. 1365–1381.

Dundurs, J., and Comninou, M., 1981, "An Educational Elasticity Problem With Friction, Part 1: Loading, and Unloading for Weak Friction," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 841–845.

Dundurs, J., and Comninou, M., 1983, "An Educational Elasticity Problem With Friction, Part 3: General Load Paths," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 77–84.

Hertz, H., 1882, "Über die Berührung fester elastischer Körper," *Journal für die reine und angewandte Mathematik,* Vol. 92, Crelle, pp. 156–171.

Iwan, W. D., 1966, "A Distributed Element Model for Hysteresis and its Steady-State Dynamic Response," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 893–900.

Jäger, J., 1992, "Frictional Impact of Elastic Bodies," Doctoral Thesis, Delft, The Netherlands.

Jäger, J., 1993, "Torsional Load-Histories of Elastic Spheres in Contact," *Contact Mechanics,* M. J. Aliabadi and C. A. Brebia, eds., Computational Mechanics Publications, Southampton, U.K., pp. 405–412.

Jäger, J., 1994a, "Analytical Solutions of Contact Impact Problems" part of "Mechanics of Contact Impact," Miloslav Okrouhlik ed., *Applied Mechanics Review,* Vol. 47, pp. 35–54.

Jäger, J., 1994b, "Oblique Impact of Similar Bodies With Circular Contact," *Acta Mechanica,* Vol. 107, pp. 101–115.

Jäger, J., 1994c, "Torsional Impact of Elastic Spheres," *Archive of Applied Mechanics,* Vol. 64, pp. 235–248.

Jäger, J., 1995, "Axi-symmetric Bodies of Equal Material in Contact Under Torsion or Shift," *Archive of Applied Mechanics,* Vol. 65, 478–487.

Jarzębowski, A., Mróz, Z., 1994, "On Slip and Memory Rules in Elastic, Friction Contact Problems," *Acta Mechanica,* Vol. 102, pp. 199–216.

Johnson, K. L., 1985, *Contact Mechanics,* Cambridge University Press, Cambridge, U.K.

Kalker, J. J., 1990, *Three-Dimensional Elastic Bodies in Rolling Contact,* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Lubkin, J. L., 1951, "The Torsion of Elastic Spheres in Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 18, pp. 183–187.

Maw, N., Barber, J. R., and Fawcett, J. N., 1976, "The Oblique Impact of Elastic Spheres," *Wear,* Vol. 38, pp. 101–114.

Mindlin, R., 1949, "Compliance of Elastic Bodies in Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 16, pp. 259–268.

Mindlin, R., and Deresiewicz, H., 1953, "Elastic Spheres in Contact Under Varying Oblique Forces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 20, pp. 327–344.

Seridi, A., and Dobry, R., 1984, "An Incremental Elastic-Plastic Model for the Force-Displacement Relation at the Contact Between Two Spheres," Research Report, Department of Civil Engineering, Rensselaer Polytechnic Inst., Troy, NY.

Tabor, D., 1981, "Friction: The Present State of Our Understanding," *Journal of Lubrication Technology,* Vol. 103, pp. 169–179.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Wrinkling of Thin, Flat, Rectangular Webs[1]

C. C. Lin
Design Engineer,
Quantum Corporation,
500 McCarthy Boulevard,
Milpitar, CA 95035
Assoc. Mem. ASME

C. D. Mote, Jr.
Vice Chancellor and FANUC Chair
in Mechanical Systems,
Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720
Fellow ASME

A web is termed wrinkled when one of the in-plane principal stresses is tensile and the other is sufficiently compressive. A criterion is derived that predicts wrinkling of isotropic, compressible rectangular webs under uniform in-plane principal stresses. The compressive stress at impending wrinkling depends on the flexural stiffness, and it equals zero in the case of a membrane. A criterion of wrinkling is also derived using isotropic, incompressible membrane theory. This criterion predicts an infinite number of wrinkle waves in a wrinkled region. With small flexural stiffness, the number of wrinkle waves becomes finite at wrinkling and it is predictable along with the shape and the size of the wrinkled region. The number of the wrinkle waves increases as the aspect ratio of the rectangular web increases, as the in-plane principal tension increases, and as the flexural stiffness decreases. Analyses of wrinkling of a rectangular web under simple shear and under uniform longitudinal stretching illustrate the above predictions.

1 Introduction

Thin, lightweight, membrane-like materials stretched in tension are found in industry as paper webs, plastic sheets, films, textiles, and other systems generally referred to as webs. The flexural stiffness of webs ranges from small to very small and wrinkling of them is a common problem. Wrinkling of the web degrades quality, increases defects, and can lead to breakage of the web. The onset of wrinkling constrains web handling so the factors causing wrinkling have to be identified and controlled.

An unsupported web normally travels under approximately uniform longitudinal tension between rollers. The web may be placed in in-plane shear when the rollers are not perfectly aligned. In that case, a compressive in-plane principal stress is created. The deflection of the web under aerodynamic transverse loading is usually large compared to its thickness, and the in-plane compression and shear that are induced by such deformation can also be significant to the impending wrinkling (Lin and Mote, 1995a). In-plane principal compressive stresses cause wrinkling in the web.

A membrane has zero flexural stiffness and wrinkles under any nonvanishing compressive stress. Reissner (1938) showed that a membrane is stressed in tension along the direction of the wrinkle waves and in compression perpendicular to the wrinkle waves. Roddeman et al. (1987) discussed three stress states resulting in taut, slack, or wrinkled membranes. A web with small flexural stiffness can support nonvanishing in-plane compressive stress of magnitude less than a critical value. For this web, four stress states are possible. If both the in-plane principal stresses are tensile, the web is taut. If both these stresses are compressive, the web is slack. If one principal stress is tensile and the other is compressive with magnitude less than a critical value, the web is smooth. If one principal stress is tensile and the other is compressive with magnitude greater than the critical value, the web is wrinkled.

In some circumstances, the boundary conditions and the aspect ratio of the web can have negligible influence on the critical compressive stress at wrinkling. In these cases, a closed-form solution for the critical compression in a rectangular web with all edges simply supported can be used to predict wrinkling in other webs, thereby generalizing the results.

This paper presents a criterion predicting the critical compressive stress at wrinkling of thin, flat, rectangular webs under uniform, in-plane, principal stresses. A second criterion, predicting wrinkling of an isotropic, incompressible membrane, is also derived. An infinite number of wrinkle waves appear in the wrinkled region of a membrane. With small flexural stiffness, the number of wrinkle waves is finite and is predictable along with the shape and the size of the wrinkled region. Wrinkling of webs under simple shear and under uniform longitudinal stretching are addressed as example problems.

2 Determination of the Critical Compression

Consider a thin, rectangular web under uniform in-plane principal tension T and compression $b \cdot T$ as shown in Fig. 1. Let θ be the angle between the direction of the in-plane principal tension and the longitudinal axis.

2.1 When $\theta = 0$ deg. When $\theta = 0$ deg, the critical compression at impending wrinkling, b_{cr}, is the minimum eigenvalue of the dimensionless equation of equilibrium,

$$-\hat{w}_{,\hat{x}\hat{x}} + b\hat{w}_{,\hat{y}\hat{y}} + \epsilon(\hat{w}_{,\hat{x}\hat{x}\hat{x}\hat{x}} + 2\hat{w}_{,\hat{x}\hat{x}\hat{y}\hat{y}} + \hat{w}_{,\hat{y}\hat{y}\hat{y}\hat{y}}) = 0, \quad (1)$$

with appropriate boundary conditions. The comma denotes partial differentiation and

$$\hat{w} = \frac{w}{h}, \quad \hat{x} = \frac{x}{L}, \quad \hat{y} = \frac{y}{L}, \quad (2a)$$

$$\epsilon = \frac{D}{TL^2}, \quad \eta = \frac{B}{L}. \quad (2b)$$

$w(x, y)$ represents the transverse displacement; $D = Eh^3/[12(1 - \nu^2)]$; E is Young's modulus; h is thickness; ν is Poisson's ratio; and b scales the compression. The material properties are uniform and constant. The stiffness ratio, ϵ, is normally a small, nondimensional ratio of the flexural stiffness to the stiffness derived from the in-plane principal tension. η is the aspect ratio

[1] The authors would like to thank the James River Corporation and the TAPPI Research Foundation for their support and interest in this work.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Sept. 30, 1994; final revision, Oct. 18, 1995. Associate Technical Editor: R. Abeyaratne.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

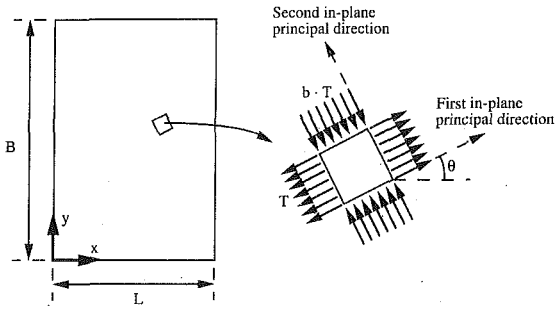


Fig. 1 A thin, rectangular web with small flexural stiffness is under uniform in-plane principal stresses. Every element in the web is subjected to a tension T and a compression $b \cdot T$ in principal directions. The angle between the direction of the in-plane principal tension and the longitudinal axis is θ.

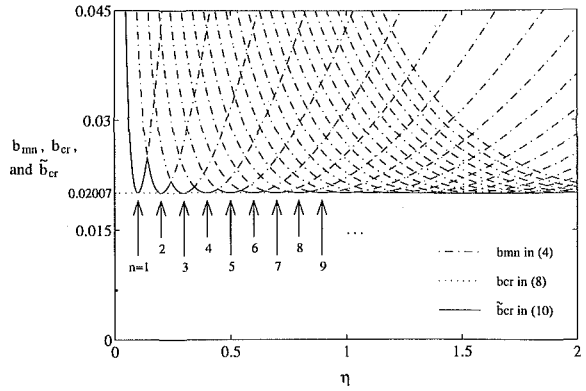


Fig. 2 b_{mn} in (4), b_{cr} in (8), and $\tilde{b}_{cr}$ in (10) for $m = 1$ and $\epsilon = 10^{-6}$

of the width to the length. For all edges simply supported, $\hat{w}$ can be represented by the form

$$\hat{w} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \hat{w}_{mn} \sin (m\pi\hat{x}) \sin \left(\frac{n}{\eta} \pi\hat{y}\right), \tag{3}$$

where $\hat{w}_{mn}$ are amplitudes of the wrinkling modes (eigenfunctions) associated with the eigenvalues

$$b_{mn} = \frac{m^2 + \epsilon\left[m^2 + \left(\frac{n}{\eta}\right)^2\right]^2 \pi^2}{\left(\frac{n}{\eta}\right)^2} \tag{4}$$

of (1) plus boundary conditions, for each integer m, $n \geq 1$. The minimum b_{mn}, associated with $m = 1$, is

$$b_{1n} = \frac{1 + \epsilon[1 + g^2]^2\pi^2}{g^2}, \tag{5}$$

where

$$g = \frac{n}{\eta}. \tag{6}$$

The only real solution to $(d/dg)(b_{1n}) = 0$ for continuous g in $(0, \infty)$ is

$$g_{cr} = \left(\frac{1 + \epsilon\pi^2}{\epsilon\pi^2}\right)^{1/4}. \tag{7}$$

Because $(d^2/dg^2)(b_{1n}) > 0$, substitution of $g = g_{cr}$ into (5) gives the minimum value of b_{1n},

$$b_{cr} = \frac{1 + \epsilon\left[1 + \left(\frac{1 + \epsilon\pi^2}{\epsilon\pi^2}\right)^{1/2}\right]^2 \pi^2}{\left(\frac{1 + \epsilon\pi^2}{\epsilon\pi^2}\right)^{1/2}}. \tag{8}$$

n is an integer and ηg_{cr} in (6) is not normally an integer. The two integers closest to ηg_{cr} are

$$n_+ = \text{Int}_+(\eta g_{cr}) \tag{9a}$$

and

$$n_- = \text{Int}_-(\eta g_{cr}). \tag{9b}$$

They are used to compute $\tilde{b}_+$ and $\tilde{b}_-$ in (5) and then

$$\tilde{b}_{cr} = \min (\tilde{b}_+, \tilde{b}_-) \tag{10}$$

is the minimum b_{mn} for all m and n.

The b_{mn} in (4), b_{cr} in (8), and $\tilde{b}_{cr}$ in (10) for $m = 1$ and $\epsilon = 10^{-5}$ are shown in Fig. 2. When $\eta > 0.3$, the difference between b_{cr} ($= 0.02007$) and $\tilde{b}_{cr}$ is less than four percent. For $\epsilon = 10^{-4}$ and $\eta > 0.3$, the difference is less than eight percent, and it is less than two percent when $\epsilon = 10^{-6}$. The differences between b_{cr} and $\tilde{b}_{cr}$ decrease as ϵ decreases. For our purposes, b_{cr} in (8) can be used to represent webs with $\epsilon < 10^{-4}$, $\eta > 0.3$ and all edges simply supported.

The b_{cr} in (8) and the $\hat{b}_{cr}$ computed for rectangular webs of different aspect ratios and boundary supports are shown in Fig. 3. The eigenvalues $\hat{b}_{cr}$ satisfy (1) and the specified boundary conditions. They are determined by Galerkin analyses. Transition between modes occurs at the points P and Q. The transitions are not easily determined from figure when $\epsilon < 10^{-6}$ or $\eta > 1$. With $\eta = 0.3$, the maximum relative error between $\hat{b}_{cr}$ and b_{cr} is less than nine percent. The difference decreases as η increases and as the mode number n_+ (or n_-) increases. Thus, it is concluded that b_{cr} in (8) applies approximately to all rectangular webs under uniform longitudinal tension T with $\eta > 0.3$ and $\epsilon < 10^{-4}$ independent of their boundary conditions.

2.2 When $\theta \neq 0$ deg. When $\theta \neq 0$ deg, the equation of equilibrium becomes

$$-T_x w_{,xx} - 2T_{xy} w_{,xy} - T_y w_{,yy} + D(w_{,xxxx} + 2w_{,xxyy} + w_{,yyyy}) = 0, \tag{11}$$

where

$$T_x = \frac{T - bT}{2} + \frac{T + bT}{2} \cos (2\theta) \tag{12a}$$

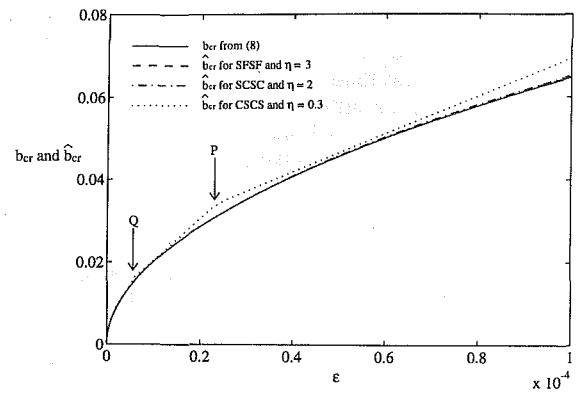


Fig. 3 b_{cr} and $\hat{b}_{cr}$ for rectangular webs of aspect ratio η. (C: clamped, S: simply supported, F: free.) (P: transition from mode $(m, n) = (1, 2)$ to $(m, n) = (1, 3)$, Q: transition from mode $(m, n) = (1, 3)$ to $(m, n) = (1, 4)$.)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

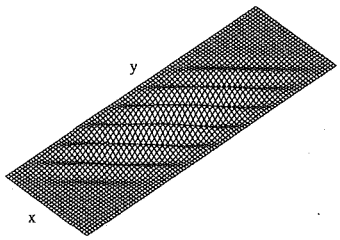


Fig. 4 **The fundamental wrinkling mode corresponding to $b_{cr} = 0.014$ of a rectangular web with $\epsilon = 10^{-5}$, $\eta = 3$, $\theta = 45$ deg, and all edges simply supported**

$$T_y = \frac{T - bT}{2} - \frac{T + bT}{2} \cos(2\theta) \qquad (12b)$$

$$T_{xy} = \frac{T + bT}{2} \sin(2\theta). \qquad (12c)$$

Here, we let 0 deg $< \theta \leq$ 45 deg because b_{cr} for $-\theta$ equals that for $+\theta$ and we can interchange the width and longitudinal axes for 45 deg $\leq \theta \leq$ 90 deg. The eigenvalue problem, formed by (11)–(12) plus simple support boundary conditions, is solved using the Galerkin method. The fundamental eigenvalue for $\epsilon = 10^{-5}$, $\eta = 3$, and $\theta = 45$ deg is $b_{cr} = 0.014$ and the mode is shown in Fig. 4. The wrinkle waves are inclined approximately 45 deg to the positive longitudinal axis. They lie parallel to the direction of the in-plane principal tension and perpendicular to the direction of the in-plane principal compression as predicted by Reissner (1938) for a membrane. Similar orientations of the wrinkle waves are observed for other ϵ, η, and θ values.

The web in Fig. 4 is partly wrinkled. No wrinkle waves appear in the lower right and upper left regions. The wrinkled region lies within a parallelogram shown in Fig. 5(a, b). Two vertices of the parallelogram are vertices of the web and two sides of the parallelogram coincide with sides of the web. θ is the angle of inclination of the remaining two sides of the parallelogram with the positive longitudinal axis. The length $\bar{L}$ of the parallelogram is $L/\cos(\theta)$ in Fig. 5(a), and $\eta L/\sin(\theta)$ in Fig. 5(b). $\bar{L}$ can be determined by

$$\bar{L} = \min\left(\frac{L}{\cos(\theta)}, \frac{\eta L}{\sin(\theta)}\right). \qquad (13)$$

The aspect ratio of the parallelogram, $\bar{\eta} = \bar{B}/\bar{L}$, is

$$\bar{\eta} = [\eta - \tan(\theta)] \cos^2(\theta) \qquad (14a)$$

in Fig. 5(a) and

$$\bar{\eta} = \left[\frac{1}{\eta} - \cot(\theta)\right] \sin^2(\theta) \qquad (14b)$$

in Fig. 5(b). The size, base length, and height of the parallelogram are equivalent to those of a rectangular web QVSU. The b_{cr} of the parallelogram differs by less than nine percent from that of QVSU when $\bar{\eta} > 0.3$ and $\bar{\epsilon} < 10^{-4}$ (see Appendix). The stiffness ratio of the web QVSU is

$$\bar{\epsilon} = \frac{D}{T\bar{L}^2} = \epsilon \frac{L^2}{\bar{L}^2}. \qquad (15)$$

Replacing ϵ in (8) by $\bar{\epsilon}$ gives b_{cr} for the web QVSU,

$$b_{cr} = \frac{1 + \bar{\epsilon}\left[1 + \left(\frac{1 + \bar{\epsilon}\pi^2}{\bar{\epsilon}\pi^2}\right)^{1/2}\right]^2 \pi^2}{\left(\frac{1 + \bar{\epsilon}\pi^2}{\bar{\epsilon}\pi^2}\right)^{1/2}}. \qquad (16)$$

Thus, b_{cr} obtained from (16) is applicable to the partly wrinkled rectangular web under uniform in-plane principal stresses with $\theta \neq 0$ deg, $\bar{\eta} > 0.3$, and $\bar{\epsilon} < 10^{-4}$. The value of b_{cr} versus θ with $\epsilon = 3 \times 10^{-5}$ and $\eta = 4$, predicted by (16), differs by less than three percent from that predicted in (11) (see Fig. 6).

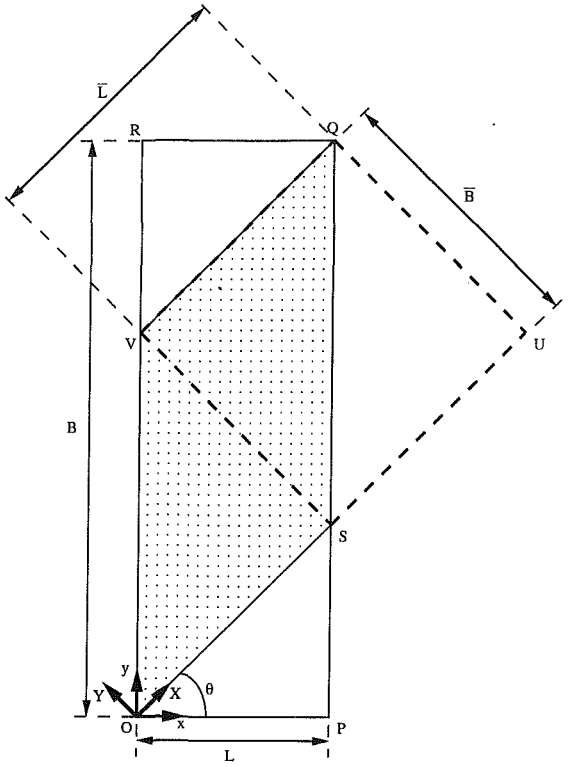


Fig. 5(a)

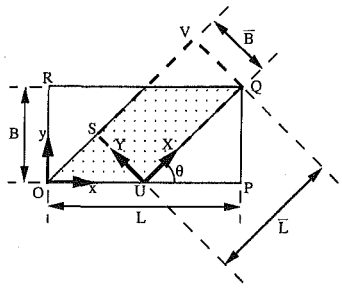


Fig. 5(b)

Fig. 5 **The wrinkled regions of a rectangular web with $\theta = 45$ deg, and (a) $\eta = 3$, and (b) $\eta = 0.5$. The shaded areas are the wrinkled regions.**

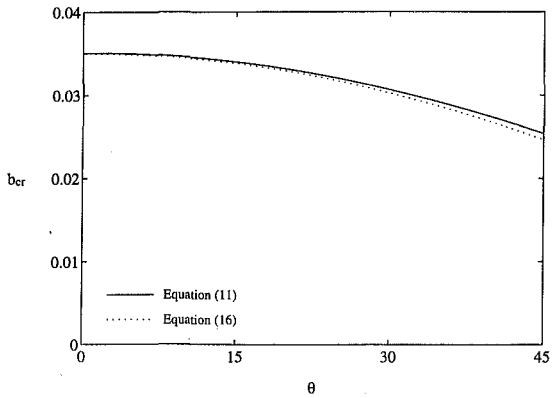


Fig. 6 **b_{cr} with $\epsilon = 3 \times 10^{-5}$ and $\eta = 4$**

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The differences are less than nine percent for webs with $\bar{\eta} > 0.3$ and $\bar{\epsilon} < 10^{-4}$.

3 Invariants of Green's Deformation Tensor

A material point in the web, initially at the position $\mathbf{s}$, deforms to a position $\mathbf{r}$. The variables in bold type are first or second-order tensors. If the web does not wrinkle under uniform in-plane stretching, $\mathbf{s}$ and $\mathbf{r}$ can be described by

$$\mathbf{s} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \tag{17}$$

and

$$\mathbf{r} = f_1(x, y)\mathbf{e}_1 + f_2(x, y)\mathbf{e}_2 + [(1 + \Delta h)z]\mathbf{e}_3. \tag{18}$$

f_1 and f_2 are linear in x and y. $\mathbf{e}_i$ are unit vectors. Δh is a dimensionless constant describing the change in thickness of the web. It is determined by the condition of plane stress.

The deformation gradient $\mathbf{F}$, defined by $d\mathbf{r} = \mathbf{F} \cdot d\mathbf{s}$, is

$$\mathbf{F} = \begin{bmatrix} f_{1,x} & f_{1,y} & 0 \\ f_{2,x} & f_{2,y} & 0 \\ 0 & 0 & 1 + \Delta h \end{bmatrix}. \tag{19}$$

Invertibility of the deformation requires $J = \det(\mathbf{F}) \neq 0$. Green's deformation tensor, $\mathbf{G} = \mathbf{F}^T \cdot \mathbf{F}$, is

$$\mathbf{G} = \begin{bmatrix} G_{11} & G_{12} & 0 \\ G_{21} & G_{22} & 0 \\ 0 & 0 & (1 + \Delta h)^2 \end{bmatrix}, \tag{20}$$

where $\mathbf{F}^T$ is $\mathbf{F}$ transpose. The elements of $\mathbf{F}$ and $\mathbf{G}$ are constant because f_1 and f_2 are linear in x and y.

Let α and β ($\alpha > \beta$) be the stretches in the in-plane principal directions and γ be the principal stretch in the $\mathbf{e}_3$ direction. α^2, β^2, and γ^2 are the eigenvalues of $\mathbf{G}$:

$$\alpha = \sqrt{\frac{I_1 + \sqrt{I_1^2 - 4I_2}}{2}} \tag{21a}$$

$$\beta = \sqrt{\frac{I_1 - \sqrt{I_1^2 - 4I_2}}{2}} \tag{21b}$$

$$\gamma = 1 + \Delta h, \tag{21c}$$

where the invariants of $\mathbf{G}$ are

$$I_1 = G_{11} + G_{22} = \alpha^2 + \beta^2 \tag{22a}$$

$$I_2 = G_{11} \times G_{22} - G_{21} \times G_{12} = \alpha^2 \times \beta^2. \tag{22b}$$

With a prescribed deformation, the angle θ, corresponding to the direction of the in-plane principal tension, can be determined from the elements of $\mathbf{G}$ (Pipkin, 1986):

$$\tan(2\theta) = \frac{2G_{12}}{G_{11} - G_{22}} \tag{23a}$$

or

$$\tan(\theta) = \frac{\alpha^2 - G_{11}}{G_{12}}. \tag{23b}$$

The direction of the wrinkle waves in the initial configuration is

$$\mathbf{v} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix}, \tag{24}$$

and in the current configuration is

$$\mathbf{u} = \mathbf{F} \cdot \mathbf{v}. \tag{25}$$

For infinitesimal deformation, $\mathbf{v} \approx \mathbf{u}$.

4 The Criterion of Wrinkling

Let H ($0 < H < 1$) be the critical stretch in the second in-plane principal direction associated with b_{cr} when the stretch in the first in-plane principal direction is $\alpha > 1$. For compressible webs, the critical stretch H depends on Poisson's ratio ν as well as b_{cr} and α. Thus, for $\beta < H$, the web wrinkles. Criteria predicting wrinkling are derived for isotropic incompressible membranes and for isotropic compressible webs.

4.1 Isotropic Incompressible Membrane. The stretch relation $\beta = \alpha^{-1/2}$ is derived for the incompressible membrane ($\alpha\beta\gamma = 1$) under simple tension ($\beta = \gamma$). The critical stretch H of an isotropic incompressible membrane is the stretch β of the membrane in simple tension so that

$$H = \alpha^{-1/2}. \tag{26}$$

$\alpha^{-1/2}$ is the natural width of the isotropic incompressible membrane in simple tension derived by Steigmann and Pipkin (1989). Compression in the second in-plane principal direction occurs and the membrane wrinkles when

$$\beta < \alpha^{-1/2}, \tag{27}$$

or, equivalently,

$$\beta^2 I_2 < 1. \tag{28}$$

Because $\alpha > 1$, the criterion of wrinkling (28) is satisfied when either

$$I_2 < 1 \tag{29a}$$

or

$$I_2 \geq 1 \quad \text{and} \quad I_1 > I_2^2 + \frac{1}{I_2}. \tag{29b}$$

4.2 Isotropic Compressible Rectangular Web. The critical stretch H of an isotropic compressible rectangular web is the infinitesimal stretch β of the web, under uniform tension T in the first in-plane principal direction and uniform compression $b_{cr} \cdot T$ in the second in-plane principal direction, so that

$$H = 1 - (\nu + b_{cr} - b_{cr}\nu^2)(\alpha - 1) + O(b_{cr}^2). \tag{30}$$

The web wrinkles when

$$\beta < 1 - (\nu + b_{cr} - b_{cr}\nu^2)(\alpha - 1). \tag{31}$$

Substitution of (22b) into the wrinkling criterion (31) gives

$$\alpha > \hat{I}_2, \tag{32}$$

where

$$\hat{I}_2 = \frac{1 + \hat{\nu} - \sqrt{(1 + \hat{\nu})^2 - 4\hat{\nu}\sqrt{I_2}}}{2\hat{\nu}}$$

$$= 1 + \frac{1}{2(1 - \hat{\nu})}(I_2 - 1) + O((I_2 - 1)^2), \tag{33}$$

$$\hat{\nu} = \nu + b_{cr} - b_{cr}\nu^2. \tag{34}$$

Multiplication of (32) by β and substitution of (22b) give the wrinkling criterion

$$\beta^2 \frac{\hat{I}_2^2}{I_2} < 1. \tag{35}$$

Because $\alpha > 1$, the criterion of wrinkling (35) is satisfied when either

$$\hat{I}_2 < 1 \tag{36a}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

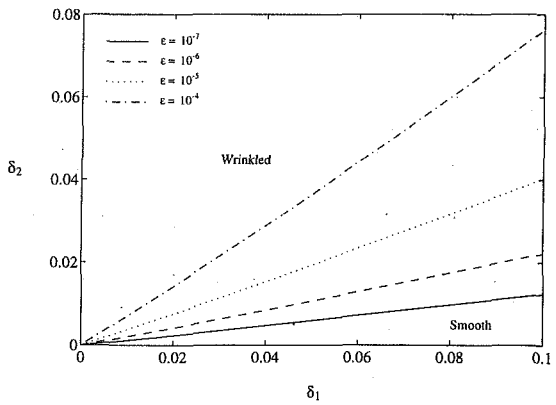


Fig. 7 Wrinkling of a rectangular web under simple shear and uniform longitudinal tension

or

$$\hat{I}_2 \geq 1 \quad \text{and} \quad I_1 > \hat{I}_2^2 + \frac{I_2}{\hat{I}_2^2}. \tag{36b}$$

For infinitesimal deformation, $I_2 \approx 1$ and $O((I_2 - 1)^2)$ in (33) can be neglected. For an isotropic, incompressible membrane ($\nu = 0.5$ and $b_{cr} = 0$), the criteria of wrinkling (36a, b) reduce to the criteria (29a, b).

5 Effects of Nonlinearly Distributed Edge Tension on Wrinkling

Linearly distributed normal tension, applied along two opposite edges or applied along all four edges of a rectangular web will not cause wrinkling. When the magnitude of the nonlinear component of the edge loading is greater than a critical value, the web wrinkles and the number of wrinkle lines in the web is finite. During manufacturing processes, an unsupported web travels under a nearly uniform, longitudinal, edge tension between the rollers. The tension can be represented as the superposition of a uniform tension and a small (perturbed), nonuniform normal edge loading. Compression can be generated in the interior region of the web when the edge tension is nonlinearly distributed. Airy stress functions can be determined in closed form for webs under arbitrarily prescribed, in-plane, edge loading. The loading at impending wrinkling, and the corresponding wrinkling shape in the web, can be predicted by the solution of an eigenvalue problem (Lin and Mote, 1995b).

6 Results and Discussion

Consider a thin, flat, rectangular web, initially occupying the region $0 < x < L$ and $-B/2 < y < B/2$, subjected to a uniform longitudinal stretch and a simple shear:

$$\mathbf{r} = [(1 + \delta_1)x]\mathbf{e}_1 + [(1 - \nu\delta_1)y + \delta_2 x]\mathbf{e}_2 + [(1 + \Delta h)z]\mathbf{e}_3. \tag{37}$$

δ_1 and δ_2 are the longitudinal and shear strains. The web is in plane stress giving $\Delta h = -\nu\delta_1$. The deformation gradient $\mathbf{F}$ and the Green's deformation tensor $\mathbf{G}$ can be determined from (19) and (20). With $\mathbf{G}$ known, θ in (23), $\bar{\epsilon}$ in (15) and b_{cr} in (16) can be determined. The invariants (22) are

$$I_1 = (1 + \delta_1)^2 + \delta_2^2 + (1 - \nu\delta_1)^2 \tag{38a}$$

and

$$I_2 = (1 + \delta_1)^2 (1 - \nu\delta_1)^2. \tag{38b}$$

For infinitesimal deformation, $I_2 > 1$ and the criterion (36b) divides the wrinkled and smooth deformations in the $\delta_1 - \delta_2$ parameter plane at $I_1 = \hat{I}_2^2 + I_2/\hat{I}_2^2$, as shown in Fig. 7. The angle of orientation of the wrinkle waves, θ, in the fundamental wrinkling mode is 2.55 deg, 4.53 deg, 8.03 deg, and 14.17 deg for the webs with stiffness ratios $\epsilon = 10^{-7}$, 10^{-6}, 10^{-5}, and 10^{-4}, respectively. For specified ϵ, δ_1 and δ_2 are linearly related at impending wrinkling. For specified δ_1, the smallest shear strain δ_2 at initial wrinkling of the web increases nonlinearly as the stiffness ratio ϵ increases. The wrinkled region is contained within a parallelogram.

An in-plane principal stress is compressive when the web is placed under simple shear and uniform longitudinal stretch. In the case of a membrane web, nonvanishing δ_2 always causes wrinkling of the entire membrane. Membrane theory underestimates the smallest δ_2 required to wrinkle the web and overestimates the size of the wrinkled area.

With $\theta = 0$ deg, the wrinkling mode is described by the product of a half-sine distribution ($m = 1$) in the longitudinal direction and a sine distribution across the width. The number of wrinkle waves N is determined by either n_+ or n_- in (9), whichever minimizes b_{mn}. Then,

$$N = \tfrac{1}{2} \times n_+ \quad \text{or} \quad N = \tfrac{1}{2} \times n_-. \tag{39a}$$

When $\theta \neq 0$ deg, the number of the wrinkle waves can be predicted by

$$N = \frac{1}{2} \times \left\{ \text{Int}_+ \left[\overline{\eta} \left(\frac{1 + \bar{\epsilon}\pi^2}{\bar{\epsilon}\pi^2} \right)^{1/4} \right] \right\} \quad \text{or}$$

$$N = \frac{1}{2} \times \left\{ \text{Int}_- \left[\overline{\eta} \left(\frac{1 + \bar{\epsilon}\pi^2}{\bar{\epsilon}\pi^2} \right)^{1/4} \right] \right\} \tag{39b}$$

whichever minimizes b_{mn}. For the rectangular web in Fig. 4, with $\epsilon = 10^{-5}$, $\eta = 3$, and $\theta = 45$ deg, the specification gives $\overline{\eta} = 1$ and $\bar{\epsilon} = 5 \times 10^{-6}$. The six wrinkle waves predicted by (39b) are observed in the numerical analysis shown in Fig. 4. The number of the wrinkle waves increases as the aspect ratio increases, as the in-plane principal tension increases, and as the flexural stiffness decreases.

For webs moving axially at constant speed v, the critical compression and the number of the wrinkle waves can be predicted by replacing the dimensional longitudinal tension T_x in (11) by $T_x - \rho v^2$, where ρ is the density of the web. Thus, the shear stress required to wrinkle the axially moving web decreases, the number of the wrinkle waves produced by shear decreases, and the angle between the wrinkle waves and the longitudinal axis increases, as the transport speed increases.

Use of the critical value b_{cr} in (16) to predict wrinkling can save substantial numerical effort required to solve the eigenvalue problems.

7 Conclusions

(1) A criterion of wrinkling (36a, b) of a thin, rectangular compressible web with small flexural stiffness is derived in terms of the invariants of Green's deformation tensor. The criterion is applicable to webs satisfying $\bar{\epsilon} < 10^{-4}$ and $\overline{\eta} > 0.3$. A criterion of wrinkling (29a, b) of incompressible membranes is also derived in terms of the invariants of Green's deformation tensor.

(2) A web wrinkles when one of the in-plane principal stresses is tensile and the other is more compressive than a critical value. This value depends on the flexural stiffness of the web and the direction and the magnitude of the in-plane principal tension. It is zero for membranes.

(3) The influence of the boundary conditions and the aspect ratio of a web on $\hat{b}_{cr}$ is less than nine percent when $\bar{\epsilon} < 10^{-4}$ and $\overline{\eta} > 0.3$.

(4) The wrinkled region of the web is a predictable parallelogram when a rectangular web wrinkles under uniform in-plane

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

principal stresses whose principal axes are not parallel to the sides of the web.

(5) The orientations of the wrinkle waves, in the web with small bending stiffness under uniform stresses, are parallel to the direction of the in-plane principal tension and perpendicular to the direction of the in-plane principal compression.

(6) Membrane theory predicts that an infinite number of wrinkle waves are produced in a wrinkled region. For web with small flexural stiffness, a finite and predictable number of wrinkle waves appear in a wrinkled region.

(7) The number of the wrinkle waves increases as the aspect ratio of the web increases, as the principal tension increases, and as the flexural stiffness decreases.

(8) Membrane theory underestimates the shear applied at the boundaries that is required to wrinkle the web and it overestimates the wrinkled area.

(9) The shear stress required to wrinkle a web under axial transport decreases, the number of the wrinkle waves produced by shear decreases, and the angle between the wrinkle waves and the longitudinal axis increases, as the transport speed increases.

(10) Linearly distributed normal tension, applied along two opposite edges or applied along all four edges of a rectangular web will not cause wrinkling. When the magnitude of the nonlinear component of the edge loading is greater than a critical value, the web wrinkles and the number of wrinkle lines in the web is finite.

References

Lin, C. C., and Mote, C. D., Jr., 1995a, "Equilibrium Displacement and Stress Distribution in a Two Dimensional, Axially Moving Web Under Transverse Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 62, pp. 772–779.

Lin, C. C., and Mote, C. D., Jr., 1995b, "Wrinkling and Stress Distribution of Rectangular Webs Under Nonlinearly Distributed Edge Loading," *Mechanics of Flexible Media in Information Storage and Processing Systems,* Proceedings of the ASME Winter Annual Meeting, San Francisco, CA.

Pipkin, A. C., 1986, "The Relaxed Energy Density for Isotropic Elastic Membranes," *IMA Journal of Applied Mathematics,* Vol. 36, pp. 85–99.

Reissner, E., 1938, "On Tension Field Theory," *Proceedings of the Fifth International Congress for Applied Mechanics,* pp. 88–92.

Roddeman, D. G., Drukker, J., Oomens, C. W. J., and Janssen, J. D., 1987, "The Wrinkling of Thin Membranes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 884–887.

Steigmann, D. J., and Pipkin, A. C., 1989, "Wrinkling of Pressurized Membranes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 624–628.

APPENDIX

Consider the parallelogram in Fig. 5 with $z = 0$ describing the midplane, and with X and Y being the in-plane coordinates. The dimensionless equation of equilibrium is

$$-\hat{w}_{,\hat{X}\hat{X}} + b\hat{w}_{,\hat{Y}\hat{Y}} + \epsilon(\hat{w}_{,\hat{X}\hat{X}\hat{X}\hat{X}} + 2\hat{w}_{,\hat{X}\hat{X}\hat{Y}\hat{Y}} + \hat{w}_{,\hat{Y}\hat{Y}\hat{Y}\hat{Y}}) = 0. \quad \text{(A1)}$$

The dimensionless parameters are

$$\hat{w} = \frac{w}{h}, \quad \hat{X} = \frac{X}{\bar{L}}, \quad \hat{Y} = \frac{Y}{\bar{L}}, \quad \text{(A2}a\text{)}$$

$$\bar{\epsilon} = \frac{D}{T\bar{L}^2}, \quad \bar{\eta} = \frac{\bar{B}}{\bar{L}}. \quad \text{(A2}b\text{)}$$

$\hat{w}$ can be represented by the form

$$\hat{w} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \sin\{m\pi[\hat{X} - \hat{Y}\tan(\theta)]\} \times \sin\left(\frac{n}{\eta}\pi\hat{Y}\right) \quad \text{(A3}a\text{)}$$

for all edges of the parallelogram simply supported in Fig. 5(a) and

$$\hat{w} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \sin\{m\pi[\hat{X} + \hat{Y}\cot(\theta)]\} \times \sin\left(\frac{n}{\eta}\pi\hat{Y}\right) \quad \text{(A3}b\text{)}$$

in Fig. 5(b). Here 0 deg $< \theta \leq$ 45 deg, a_{mn} are constants, and $m, n = 1, 2, 3, \ldots$. The minimum eigenvalue b_{cr}, from (A1) plus appropriate boundary conditions, differs from b_{cr} in (16) by less than nine percent when $\bar{\eta} > 0.3$ and $\bar{\epsilon} < 10^{-4}$. The difference decreases as $\bar{\eta}$ increases and as $\bar{\epsilon}$ decreases.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

An Approximate Solution for the Compression of a Bonded Thin Annular Disk

Yun Ling
Technology Group,
AMP, Inc.
P. O. Box 3608, M.S. 106-13,
Harrisburg, PA 17105-3608
Mem. ASME

An approximate solution for the compression of a bonded thin annular disk is presented based on the so-called Perturbation-Ritz Method. The solution is essentially the outer expansion of the problem, with the unknown constants determined by the Ritz Method minimizing the potential energy. It is valid throughout the thin disk except in the boundary layers, which are confined to very narrow regions near the lateral surfaces. The solution asymptotically approaches the exact one as the disk thickness reduces towards zero. Both the incompressible and compressible cases are discussed. The relationships of the stiffness and stress distributions with the key parameters such as the shape factor, the radius ratio, and Poisson's ratio are investigated. A better understanding of the compression of a bonded thin annular disk is achieved.

1 Introduction

The compression of a linearly elastic circular disk can be classified into the *end problem of a cylinder,* which has been subjected to numerous investigations (e.g., Swan, 1968). There are two different approaches to this problem. One is the engineering approach (e.g., Gent and Lindley, 1959; Chalhoub and Kelly, 1991; Ling et al., 1992), developed mainly for the stiffness calculation of thin rubber disks. This approach, however, inherently involves some intuitive assumptions and lacks rigor. The other is the rigorous approach (e.g., Benthem and Minderhoud, 1972; Robert and Keer, 1989; Ling et al., 1994), which uses mainly the eigenfunction expansion method. Unfortunately, the solutions based on this approach are difficult to use in practice due to the tedious mathematical work and the slow convergence of the series.

The poor convergence arises from the fact that there are stress singularities on the circumferences of the disk. The work by Benthem and Minderhoud (1972), Gent et al. (1974), and Ling et al. (1994) suggested that the "strength" of the singularities reduces quickly as the disk thickness decreases, and for a thin disk, the effect of the stress singularities is confined to a very narrow region near the circumferences. In fact, the problem under consideration is a boundary layer problem, as suggested by the studies of the related plate problem (e.g., Reiss, 1962; Gregory and Wan, 1984). The boundary layers lie near the circumferences. It is well known that a boundary layer problem may be described by the inner and outer solutions. The inner solution characterizes the behavior near the circumferences. It decays quickly away from the circumferences for a thin disk. The outer solution should then be a good approximation throughout the disk except near the circumferences.

The present study is motivated by the calculation of the stiffness and stress distributions in bonded thin *annular* rubber disks, which are frequently used as vibration isolators and sealing components. Figure 1 is the illustration of an annular disk. The purpose is to develop an *approximate* solution for the compression of the disks, eliminating the intuitive assumptions used in the engineering approach.

Historically, little attention has been paid to annular disks; most of the studies are on plane-strain rectangular strips and solid circular disks, which are special cases of annular disks. And there is still no known rigorous solution to this problem available. The engineering approach proposed by Ling et al. (1995) was in reasonable agreement with the numerical solution, but there was no rigorous justification for their assumption of the displacement fields.

The confinement to *thin* disks, by which we mean that the parameter $\epsilon = L'/(r_2' - r_1') \ll 1$ (where $2L'$ is the disk height and r_1' and r_2' are, respectively, the inner and outer radii, see Fig. 1), is justified by the fact that in practice, almost all blocks in compression can be described as *thin.* Then, the outer expansion may be used as an approximate solution to the problem. It is well known that there are unknown constants remaining in the outer solution for a boundary layer problem. If the inner solution is available, these constants could be obtained by matching the inner and outer solutions. However, since we have no intention of finding the inner solution, how should these constants be determined without using the inner solution? Though Saint-Venant's Principle may be used here, in a similar way as in the plate theory, we propose to use the Ritz Method (Fung, 1965), i.e., the constants are determined by minimizing the strain energy. This method is very similar to the Perturbation-Galerkin Method proposed by Geer and Andersen (1989, 1990). They suggested using the perturbation series as the trial functions, with their coefficients determined by the Galerkin method, and found that the solutions are much more accurate than the truncated perturbation series.

The discussion begins with the mathematical formulation of the problem. Then the so-called Perturbation-Ritz Method is described. The application of the method to the simple incompressible case is then presented; the solution is compared with the numerical results of a finite element analysis. Finally, the general compressible case is discussed.

2 Mathematical Formulation

A circular annular disk, occupying the region $r_1' \le r' \le r_2'$ and $-L' \le z' \le +L'$ (see Fig. 1), is considered. The disk is bonded to two horizontal rigid plates such that lateral displacements at the two plane ends are completely constrained. The disk is compressed over a distance $2\delta'$ by the rigid plates, which remain horizontal. The problem is obviously axisymme-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 9, 1995; final revision, Feb. 8, 1996. Associate Technical Editor: W. K. Liu.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

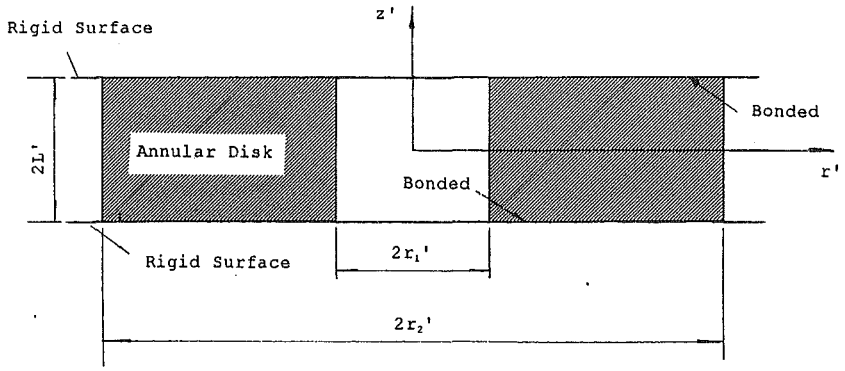


Fig. 1 Illustration of an annular disk

tric. The radial and axial displacements are, respectively, u' and w'. The normal stresses are σ_r', σ_θ', and σ_z', and the nonzero shear stress is τ_{rz}'. Nondimensional quantities will be used, defined as

$$\epsilon = \frac{L'}{r_2' - r_1'}, \quad r = \frac{r'}{r_2' - r_1'}, \quad z = \frac{z'}{L'}, \quad e = \frac{\delta'}{L'},$$

$$u = \frac{u'\epsilon}{\delta'}, \quad w = \frac{w'}{\delta'}, \quad \tau_{rz} = \frac{\tau_{rz}'\epsilon}{2G'e},$$

$$\sigma_r = \frac{\sigma_r'\epsilon^2}{2G'e}, \quad \sigma_z = \frac{\sigma_z'\epsilon^2}{2G'e}, \quad \text{etc.,} \tag{1}$$

where G' is the shear modulus. Note that the prime denotes dimensional quantities. Using the Love strain function ϕ, normalized by $2G'eL'^3/\epsilon^2$, the equilibrium equations and the deformation compatibility condition are reduced to one biharmonic equation (e.g., Benthem and Minderhoud, 1972):

$$\nabla^2\nabla^2\phi = 0, \tag{2}$$

where ∇^2 is the Laplacian operator in a slightly different form due to the normalization:

$$\nabla^2 = \frac{\partial^2}{\partial z^2} + \epsilon^2\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right). \tag{3}$$

The stresses can be expressed in terms of ϕ:

$$\sigma_r = \frac{\partial}{\partial z}\left\{\nu\nabla^2\phi - \epsilon^2\frac{\partial^2\phi}{\partial r^2}\right\}, \tag{4a}$$

$$\sigma_\theta = \frac{\partial}{\partial z}\left\{\nu\nabla^2\phi - \frac{\epsilon^2}{r}\frac{\partial\phi}{\partial r}\right\}, \tag{4b}$$

$$\sigma_z = \frac{\partial}{\partial z}\left\{(2-\nu)\nabla^2\phi - \frac{\partial^2\phi}{\partial z^2}\right\}, \tag{4c}$$

$$\tau_{rz} = \frac{\partial}{\partial r}\left\{(1-\nu)\nabla^2\phi - \frac{\partial^2\phi}{\partial z^2}\right\}, \tag{4d}$$

where ν is Poisson's ratio. The displacements are then

$$u = -\frac{\partial^2\phi}{\partial r\partial z}, \tag{5a}$$

$$w = \frac{1}{\epsilon^2}\left\{2(1-\nu)\nabla^2\phi - \frac{\partial^2\phi}{\partial z^2}\right\}. \tag{5b}$$

The boundary conditions at the plane ends can be written as

$$z = +1: \quad u = 0, \quad w = -1, \tag{6a}$$

$$z = -1: \quad u = 0, \quad w = +1. \tag{6b}$$

The lateral surfaces are stress-free:

$$r = r_1 = \frac{\beta}{1-\beta}: \quad \sigma_r = 0, \quad \tau_{rz} = 0, \tag{7a}$$

$$r = r_2 = \frac{1}{1-\beta}: \quad \sigma_r = 0, \quad \tau_{rz} = 0, \tag{7b}$$

where $\beta = r_1'/r_2'$ is referred to as the *radius ratio.*

3 Perturbation-Ritz Method

The problem is now confined to a thin annular disk, for which $\epsilon \ll 1$. Then the outer solution for the strain function ϕ may be expanded into a power series of ϵ^2:

$$\phi = \sum_{j=0}^{\infty} \phi_{2j}\epsilon^{2j}, \tag{8}$$

where ϕ_{2j} are functions to be solved, independent of ϵ. We also assume that the material is only slightly compressible, i.e., $1 - 2\nu$ is small and we can set $1 - 2\nu = \kappa\epsilon^2$, where κ is a constant. Then, the corresponding displacements u and w are

$$u = \sum_{j=0}^{\infty} u_{2j}\epsilon^{2j}, \quad w = \sum_{j=0}^{\infty} w_{2j}\epsilon^{2j}, \tag{9}$$

where

$$u_{2j} = -\frac{\partial^2\phi_{2j}}{\partial r\partial z}, \tag{10a}$$

$$w_{2j} = \left(\kappa\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\phi_{2j} + \kappa\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\phi_{2(j-1)}. \tag{10b}$$

The differential equation for ϕ_{2j} is obtained from Eq. (2) by collecting terms containing ϵ^{2j}:

$$\frac{\partial^4\phi_{2j}}{\partial z^4} + 2\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\frac{\partial^2\phi_{2(j-1)}}{\partial z^2} + \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)^2\phi_{2(j-2)} = 0, \tag{11}$$

where $j = 0, 1, 2, \ldots$, and $\phi_{-2} = \phi_{-4} \equiv 0$.

Obviously, the outer solution has to satisfy the displacement boundary conditions at the plane ends:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$z = \pm 1: \quad u_{2j} = 0, \quad w_{2j} = \begin{cases} \mp 1, & j = 0 \\ 0, & j > 0 \end{cases}. \tag{12}$$

However, the solutions defined by Eqs. (11) and (12) are not expected to satisfy the traction-free boundary conditions on the lateral surfaces. The traction-free conditions may only be approximately satisfied by the outer solution. One way to achieve this is to apply Saint-Venant's Principle, requiring that the unknown constants in the outer solution be chosen such that

$$\int_{-1}^{+1} \sigma_r(r_1, z)dz = \int_{-1}^{+1} \sigma_r(r_2, z)dz = 0, \tag{13a}$$

$$\int_{-1}^{+1} \tau_{rz}(r_1, z)dz = \int_{-1}^{+1} \tau_{rz}(r_2, z)dz = 0. \tag{13b}$$

Since conditions (13) are statically equivalent to conditions (7), using them should affect the stresses only in the vicinity of the lateral surfaces, according to Saint-Venant's Principle.

It is realized, however, that for an elasticity problem, one is trying to determine stresses and strains for given boundary conditions such that the system has a minimum potential energy. Thus, if the trial functions are given, the best way to determine the coefficients may be to make the potential energy a minimum. This is the Ritz Method (e.g., Fung, 1965). The outer solutions for the strain function and the displacements truncated at order ϵ^{2n} can be expressed as

$$\phi \approx \sum_{j=0}^{n} \phi_{2j}\epsilon^{2j} = f_0 + \sum_{i=1}^{m} a_i f_i, \tag{14a}$$

$$u \approx \sum_{j=0}^{n} u_{2j}\epsilon^{2j} = g_0 + \sum_{i=1}^{m} a_i g_i, \tag{14b}$$

$$w \approx \sum_{j=0}^{n} w_{2j}\epsilon^{2j} = h_0 + \sum_{i=1}^{m} a_i h_i, \tag{14c}$$

where a_i ($i = 1$ to m) are unknown constants or the coefficients, and f_i, g_i, and h_i are *known* functions obtained by rearranging the perturbation solutions. Note that although f_i, g_i, and h_i may be called the trial functions, they were obtained by the regular perturbation method, not by trials. The coefficients a_i here are allowed to be functions of ϵ as in the Perturbation-Galerkin method suggested by Geer and Andersen (1989, 1990). Their work has indicated that the combination of the perturbation solution and the energy method often leads to exceptionally good results.

The normalized strain energy reads

$$U = \tfrac{1}{2}\int_{-1}^{+1}\int_{r_1}^{r_2} (\sigma_r\epsilon_r + \sigma_\theta\epsilon_\theta + \sigma_z\epsilon_z + \tau_{rz}\gamma_{rz})rdrdz, \tag{15}$$

where the normalized strains ϵ_r, ϵ_θ, ϵ_z, and γ_{rz} are defined as

$$\epsilon_r = \frac{\partial u}{\partial r}, \quad \epsilon_\theta = \frac{u}{r}, \quad \epsilon_z = \frac{\partial w}{\partial z}, \quad \gamma_{rz} = \frac{\partial u}{\partial z} + \epsilon^2\frac{\partial w}{\partial r}. \tag{16}$$

Note that the stresses in Eq. (15) are evaluated by substituting the approximate ϕ into Eqs. (4), while the strains are calculated using the displacements given in Eqs. (14*b*) and (14*c*), which satisfy all the displacement boundary conditions and thus meet the requirement of using the Ritz Method. The strain energy U then becomes the function of the coefficients a_i. The Ritz Method suggests that a_i be determined such that the strain energy is minimized, i.e.:

$$\frac{\partial U}{\partial a_i} = 0, \quad i = 1, 2, \ldots, m. \tag{17}$$

One of the main parameters sought in the calculation is the stiffness, the ratio of the total force F' to the total compression $2\delta'$. The force F' is given by

$$F' = -2\pi\int_{r_1}^{r_2} \sigma_z' r'dr' = \frac{-4\delta'}{(1+\nu)(r_1+r_2)\epsilon^2}\left(\frac{E'A'}{2L'}\right)\int_{r_1}^{r_2}\sigma_z rdr, \tag{18a}$$

where $E' = 2(1+\nu)G'$ is the elastic modulus and A' is the cross-sectional area (perpendicular to the z-axis). Note that F' should be independent of the coordinate z, but Eq. (18*a*) may not be able to fulfill this requirement since the solution sought is approximate. Thus, we choose to calculate F' by using the stress σ_z averaged over $-1 \le z \le +1$ in Eq. (18*a*). Then the stiffness, normalized by $E'A'\delta'/(2L')$, becomes

$$K = \frac{-2}{(1+\nu)(r_1+r_2)\epsilon^2}\int_{r_1}^{r_2}\left[\frac{1}{2}\int_{-1}^{+1}\sigma_z dz\right]rdr. \tag{18b}$$

4 Solution for the Incompressible Case

Rubber is frequently approximated as an incompressible material, whose Poisson's ratio equals $\frac{1}{2}$, or $\kappa = 0$. ϕ_0 can be solved from Eqs. (11) for $j = 0$ as

$$\phi_0 = A_0(r) + A_1(r)z + A_2(r)z^2 + A_3(r)z^3, \tag{19a}$$

where A_i ($i = 0$ to 3) are functions of r. Using boundary conditions (12), we find that

$$A_0(r) = A_2(r) = 0, \quad A_3(r) = -\frac{A_1(r)}{3}, \tag{20}$$

where $A_1(r)$ is determined by

$$\frac{d^2A_1}{dr^2} + \frac{1}{r}\frac{dA_1}{dr} = -\frac{3}{2}. \tag{21}$$

The solution of Eq. (21) is

$$A_1(r) = c_1 \ln r + c_2 - \tfrac{3}{8}r^2 \tag{22}$$

where c_1 and c_2 are constants. Thus, ϕ_0 can be expressed as

$$\phi_0 = \left(c_1 \ln r + c_2 - \frac{3}{8}r^2\right)\left(z - \frac{z^3}{3}\right). \tag{19b}$$

The function ϕ_2 is obtained by substituting the expression for ϕ_0 into Eqs. (11) for $j = 1$:

$$\phi_2 = B_1(r)\left(z - \frac{z^3}{3}\right) - \frac{z^5}{20}, \tag{23a}$$

where $B_1(r)$ is a function of r, to be determined by the condition $w_2(r, 1) = 0$:

$$\frac{d^2B_1}{dr^2} + \frac{1}{r}\frac{dB_1}{dr} = 0. \tag{24}$$

The solution for Eq. (24) is

$$B_1 = c_3 \ln r + c_4, \tag{25}$$

where c_3 and c_4 are constants. Thus, the expression for ϕ_2 is

$$\phi_2 = (c_3 \ln r + c_4)\left(z - \frac{z^3}{3}\right) - \frac{z^5}{20}. \tag{23b}$$

It is found that the subsequent solutions of $\phi_{2j}(j > 1)$ can be chosen as zero. Then the strain function can be expressed as

$$\phi \approx \phi_0 + \epsilon^2\phi_2 = f_0 + a_1 f_1 + a_2 f_2, \tag{26}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$f_0 = -\left[\frac{3}{8}\left(z - \frac{z^3}{3}\right) + \epsilon^2 \frac{z^5}{20}\right],$$

$$f_1 = \left(z - \frac{z^3}{3}\right) \ln r, \quad f_2 = z - \frac{z^3}{3}. \tag{27}$$

It is easy to verify that solution (26) exactly satisfies the governing Eq. (2) and the displacement boundary conditions at the plane ends. The displacements are obtained using Eqs. (14*b*) and (14*c*):

$$u \approx g_0 + a_1 g_1, \quad w \approx h_0, \tag{28a}$$

where

$$g_0 = \frac{3}{4} r(1 - z^2), \quad g_1 = -\frac{1 - z^2}{r},$$

$$h_0 = -\frac{3}{2}\left(z - \frac{z^3}{3}\right). \tag{28b}$$

Equations (28) suggest that away from the boundary layer regions, a plane originally perpendicular to the z-axis shall remain plane and a line element initially vertical shall become a parabola after compression. This is exactly the assumptions used by Gent and Lindley (1959) for the plane-strain rectangular strips and solid cylinders. It is also seen from Eqs. (28) that $u \equiv 0$ when $r = r_n = (2/\sqrt{3})\sqrt{a_1}$. We may refer to the surface defined by $r = r_n$ as the *neutral surface*. The material expands in opposite directions on the two sides of the neutral surface. The shear stress is expected to be zero at $r = r_n$ and to change signs across the neutral surface. We can also verify that the displacements obtained satisfy the incompressibility condition:

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z} = 0. \tag{29}$$

The stresses are obtained by substituting Eq. (26) into Eqs. (4):

$$\sigma_r = \frac{3}{8}(r^2 - 4\epsilon^2 z^2) + a_1\left[\frac{\epsilon^2}{r^2}(1 - z^2) - \ln r\right] - a_2, \tag{30a}$$

$$\sigma_\theta = \frac{3}{8}(r^2 - 4\epsilon^2 z^2) - a_1\left[\frac{\epsilon^2}{r^2}(1 - z^2) + \ln r\right] - a_2, \tag{30b}$$

$$\sigma_z = \frac{3}{8}(r^2 + 2\epsilon^2 z^2 - 6\epsilon^2) - a_1 \ln r - a_2, \tag{30c}$$

$$\tau_{rz} = \left(\frac{a_1}{r} - \frac{3}{4} r\right) z. \tag{30d}$$

To determine the coefficients a_i, we first use Saint-Venant's Principle. The condition (13*b*) is automatically satisfied; the coefficients given by condition (13*a*) are

$$a_1 = \frac{9}{8} \frac{r_1^2 r_2^2 (r_1 + r_2)}{3r_1^2 r_2^2 \ln (r_2/r_1) + 2\epsilon^2 (r_1 + r_2)}, \tag{31a}$$

$$a_2 = \frac{3}{8} r_1^2 - \frac{\epsilon^2}{2} + \left[\frac{2}{3}\frac{\epsilon^2}{r_1^2} - \ln r_1\right] a_1. \tag{31b}$$

If the Ritz method is used, we obtain from Eq. (17)

$$a_1 = \frac{9}{8} \frac{r_1^2 r_2^2 (r_1 + r_2)}{\frac{12}{5}\epsilon^2 (r_1 + r_2) + 3r_1^2 r_2^2 \ln (r_2/r_1)}. \tag{32a}$$

Note that a_2 cannot be determined from Eq. (17) since the material is incompressible. However, since the solution obtained satisfies all but the traction-free conditions on the lateral surfaces, we may require that

$$\int_{-1}^{+1} (\sigma_r \delta u + \tau_{rz} \delta w)\big|_{r=r_1}^{r=r_2} dz = 0. \tag{33a}$$

Note that here the variation operation δ is with respect to a_1 and a_2. Since w is independent of a_1 and a_2, and u depends only on a_1, Eq. (33*a*) is reduced to

$$\int_{-1}^{+1} \sigma_r \frac{\partial u}{\partial a_1}\bigg|_{r=r_1}^{r=r_2} dz = 0. \tag{33b}$$

This equation gives

$$a_2 = -\frac{3}{8} r_1 r_2 + (r_1 \ln r_2 - r_2 \ln r_1) a_1 + \epsilon^2\left(\frac{4}{5} a_1 \frac{r_2^3 - r_1^3}{r_1^2 r_2^2} - \frac{3}{10}\right). \tag{32b}$$

The stiffness K is then found as

$$K = \frac{4}{3\epsilon^2}\left\{\frac{2\epsilon^2 + a_2}{2} + \frac{a_1}{4}\left[\frac{2(r_2^2 \ln r_2 - r_1^2 \ln r_1)}{r_1 + r_2} - 1\right] - \frac{3}{32}(r_2^2 + r_1^2)\right\}. \tag{34}$$

To assess the results presented, a numerical solution is sought for comparison. The finite element method is used to obtain the numerical solution. Axisymmetric elements are used to model the disk and the meshes are illustrated in Fig. 2. The numerical results from the finite element calculation are deemed reliably accurate for comparison. Figure 3 is the error e_K for the stiffness K; error e_K is defined as $|(K_a - K_n)/K_n|$, where K_a represents the analytical solution while K_n the numerical one. It can be seen that the error for the solution based on Saint-Venant's Principle is much greater than that based on the Ritz Method, except when ϵ is small. The coefficients determined by the two methods can be shown to be identical as $\epsilon \to 0$. In fact, both solutions asymptotically approach the exact solution as $\epsilon \to 0$, but the one based on the Ritz Method converges at a much faster rate. For example, for $\epsilon = .125$, the Ritz Method gives an error of about five percent while the error obtained using Saint-Venant's Principle is about 40 percent. Thus, in the rest of the paper, no further discussion will be given to Saint-Venant's Principle, since the solution based on it is less accurate. The comparisons for normal and shear stress distributions at the plane ends are shown in Figs. 4 and 5, respectively. Clearly, the solution is not valid near the circumferences, but the invalid zones are very narrow for small ϵ. The overall correlation for small ϵ, say $\epsilon \leq .10$, between the analytical and numerical solutions is very good.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

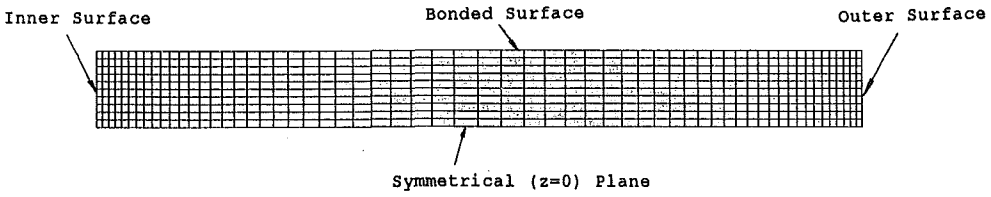


Fig. 2 Finite element mesh of an annular disk

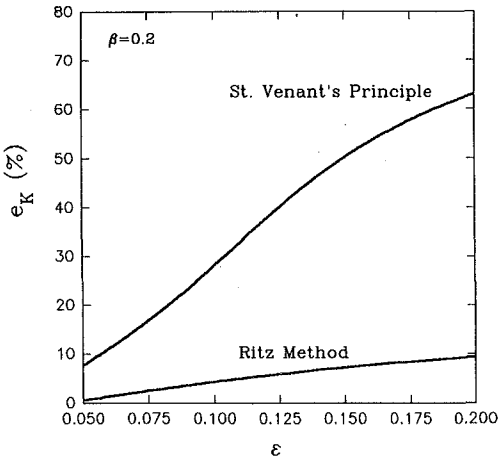


Fig. 3 Error for the analytical solution compared with the numerical one

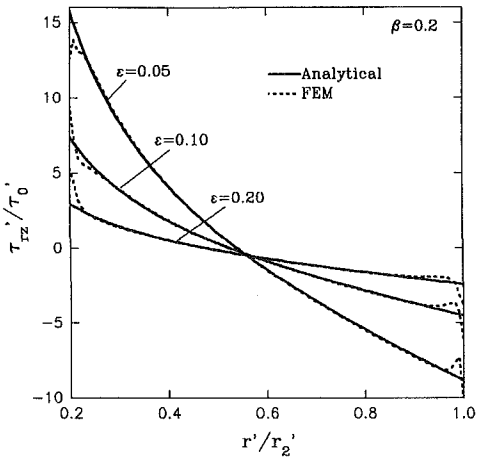


Fig. 5 Comparison of shear stresses τ_{rz}' at the plane end calculated by the Perturbation-Ritz Method and the finite element method. The normalization constant is $\tau_0' = \sigma_0' = E'e$.

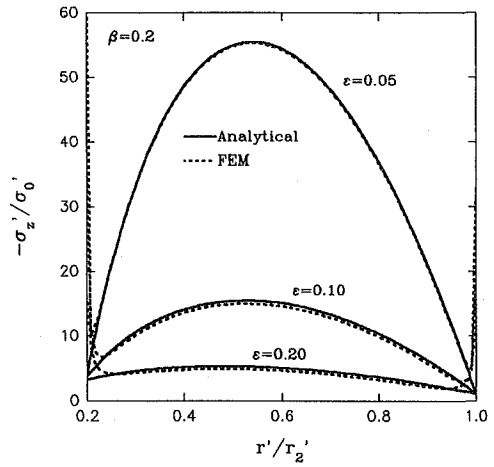


Fig. 4 Comparison of normal stresses σ_z' at the plane end calculated by the Perturbation-Ritz Method and the finite element method. The normalization constant is $\sigma_0' = E'e$.

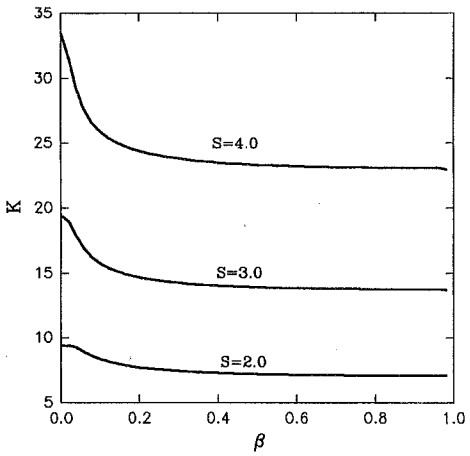


Fig. 6 Change of stiffness K with shape factor S and radius ratio β

In rubber spring design, a frequently used design parameter is the *shape factor S* defined as the *ratio of one loaded area to the force free area.* The shape factor for the annular circular disk is related to ϵ by $S = (r_2' - r_1')/(4L') = 1/(4\epsilon)$. The relationship between the stiffness K and the parameters β and S is illustrated in Fig. 6. For a disk with large shape factor (*thin* disk) there is a significant drop in stiffness as β varies from zero. The case $\beta = 0$ corresponds to the solid disk. The stiffness in this case is reduced to

$$K = \tfrac{17}{15} + 2S^2. \tag{35}$$

Note that the stiffness for a solid disk derived by Gent and Lindley (1959) is $1 + 2S^2$, which is also the result of Eq. (34) if Saint-Venant's Principle is used to determine the coefficients. As β increases, the solution quickly approaches that for the plane-strain rectangular strip, which can be obtained from Eq. (34) by letting $\beta \to 1$:

$$K = \tfrac{23}{15} + \tfrac{4}{3} S^2. \tag{36}$$

The solution given by Gent and Lindley for this case is $K = (\frac{4}{3})(1 + S^2)$.

One interesting observation is that for a very thin disk, the stiffness reduces dramatically near $\beta = 0$. The presence of a small hole in the center of a disk creates a stress-free surface allowing lateral expansion of the material and forces the "pressure" to drop, as illustrated in Fig. 7. A slight increase of β near $\beta = 0$ causes significant reduction of stress σ_z. The stress σ_z reaches a maximum at $r = r_n$ since $d\sigma_z/dr = 0$ there. The location of the neutral surface is illustrated in Fig. 8. For $\beta = 0$ or a solid disk, the neutral surface is reduced to the center

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

line $r_n = 0$. When β is near 1 (in this case, an annular disk can be considered a plane-strain rectangular strip), we find that $r_n - r_1$ approaches $\frac{1}{2}$, i.e., the neutral surface for the strip lies at the center line of the width. This is also expected.

5 Solution for the Compressible Case

For the compressible case, Eqs. (19a) and (20) are still valid. But the equation for function A_1 is

$$\frac{d^2A_1}{dr^2} + \frac{1}{r}\frac{dA_1}{dr} - 3\kappa A_1 = -\frac{3}{2}. \quad (37)$$

Its solution can be expressed in terms of Bessel Functions:

$$A_1(r) = \frac{1}{2\kappa} + c_1 I_0(\alpha r) + c_2 K_0(\alpha r), \quad (38)$$

where $\alpha = \sqrt{3\kappa}$, and I_0 and K_0 are, respectively, the Modified Bessel and Hankel Functions of zero order, and c_1 and c_2 are constants. Thus ϕ_0 is expressed as

$$\phi_0 = \left[\frac{1}{2\kappa} + c_1 I_0(\alpha r) + c_2 K_0(\alpha r)\right]\left(z - \frac{z^3}{3}\right). \quad (39)$$

Though, technically, the solution for ϕ_2 can be found, we stop at the leading term ϕ_0 to make the algebraic work manageable. The strain function is approximated as

$$\phi \approx \phi_0 = \left[\frac{1}{2\kappa} + a_1 I_0(\alpha r) + a_2 K_0(\alpha r)\right]\left(z - \frac{z^3}{3}\right). \quad (40)$$

The displacements are then

$$u \approx u_0 = -\alpha[a_1 I_1(\alpha r) - a_2 K_1(\alpha r)](1 - z^2), \quad (41a)$$

$$w \approx w_0 = -z + \kappa[a_1 I_0(\alpha r) + a_2 K_0(\alpha r)](1 - z^2)z, \quad (41b)$$

where I_1 and K_1 are the Modified Bessel and Hankel Functions of the first order. An immediate observation from Equations (41) is that when the material is compressible, although a vertical line still becomes a parabola after compression, a plane originally perpendicular to the z-axis is no longer a plane except at the plane ends $z = \pm 1$ and the symmetrical plane $z = 0$. There still exists a neutral surface defined by $r = r_n$, where r_n is determined from

$$a_1 I_1(\alpha r_n) - a_2 K_1(\alpha r_n) = 0. \quad (42)$$

The stresses are obtained by substituting Eq. (40) into Eqs. (4):

$$\begin{Bmatrix} \sigma_r \\ \sigma_\theta \\ \sigma_z \\ \tau_{rz} \end{Bmatrix} \approx -\frac{\nu}{\kappa}\begin{Bmatrix} 1 \\ 1 \\ \dfrac{1-\nu}{\nu} \\ 0 \end{Bmatrix} - \begin{bmatrix} I_0(\alpha r)\alpha_0 - I_1(\alpha r)\alpha_1/r & K_0(\alpha r)\alpha_0 + K_1(\alpha r)\alpha_1/r \\ I_0(\alpha r)\beta_0 + I_1(\alpha r)\beta_1/r & K_0(\alpha r)\beta_0 - K_1(\alpha r)\beta_1/r \\ I_0(\alpha r)\gamma_0 & K_0(\alpha r)\gamma_0 \\ -I_1(\alpha r)\delta_1 & K_1(\alpha r)\delta_1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad (43)$$

in which $\alpha_1 = \beta_1 = \alpha\epsilon^2(1 - z^2)$ and

$$\alpha_0 = 3(1 + 2\nu^2)(1 - z^2) + \nu(9z^2 - 7),$$

$$\beta_0 = 6\nu^2(1 - z^2) - \nu(1 - 3z^2),$$

$$\gamma_0 = -6\nu^2(1 - z^2) + z^2(6 - 15\nu) + 13\nu - 4,$$

$$\delta_1 = \alpha z[3 - 7\nu + 6\nu^2 + z^2(3\nu - 2\nu^2 - 1)]. \quad (44)$$

The expression for the stiffness can be found from Eq. (18b):

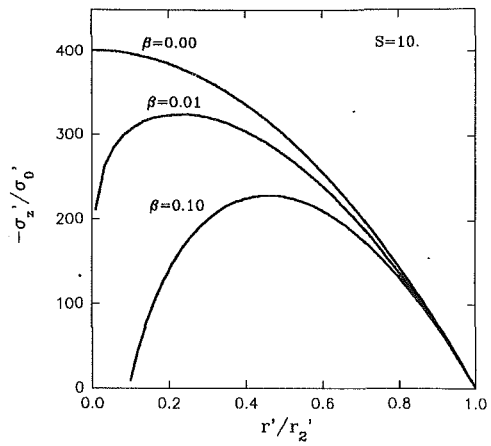


Fig. 7 Dramatic change of distributions of stress σ_z' at the plane end near $\beta = 0$ for a thin disk ($S = 10$). The normalization constant is $\sigma_0' = E'e$.

$$K = \frac{2}{(1+\nu)(r_1 + r_2)\epsilon^2}\left[\frac{1-\nu}{2\kappa}(r_2 + r_1) + \frac{2}{\alpha}(4\nu - 1 - 2\nu^2)(a_1\Omega_I - a_2\Omega_K)\right], \quad (45)$$

where $\Omega_I = r_2 I_1(\alpha r_2) - r_1 I_1(\alpha r_1)$ and $\Omega_K = r_2 K_1(\alpha r_2) - r_1 K_1(\alpha r_1)$. The coefficients a_i ($i = 1, 2$) can be obtained by substituting the displacements (41) and stresses (43) into Eq. (17). The explicit expressions for a_i are too lengthy to give here, but their numerical calculation is straightforward.

The relationships between the stiffness K and the shape factor S, the radius ratio β, and Poisson's ratio ν are illustrated in Figs. 9 and 10. The most striking observation is that for a thin disk, slight compressibility reduces the stiffness dramatically. For example, when Poisson's ratio changes from $\frac{1}{2}$ to .499, the stiffness K reduces from about 141 to 71 (for $S = 10$). Thus for a thin disk in compression, even if the rubber is almost incompressible, its compressibility should still be accounted for. This has also been suggested by Chalhoub and Kelly (1991) in their work on compressible infinite strips.

Another interesting observation is that as the material becomes more compressible, the decrease in stiffness near $\beta = 0$ becomes less significant. For example, when Poisson's ratio is .490, the stiffness becomes almost independent of the radius ratio, i.e., the stiffness for a solid disk and for a plane-strain rectangular strip are almost the same (for the same shape factor). The comparison with the finite element results in Fig. 9 also shows that although the agreement between the analytical and numerical solutions is very good, the analytical solution becomes less accurate as Poisson's ratio decreases. This is understandable since in seeking the outer expansion, the perturbation is around $\nu = \frac{1}{2}$. Figures 11 and 12 show that compressibility has a significant effect on the stress distributions.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

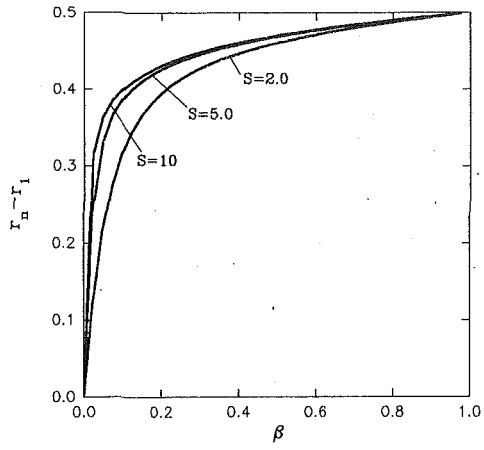


Fig. 8 Location of the neutral surface

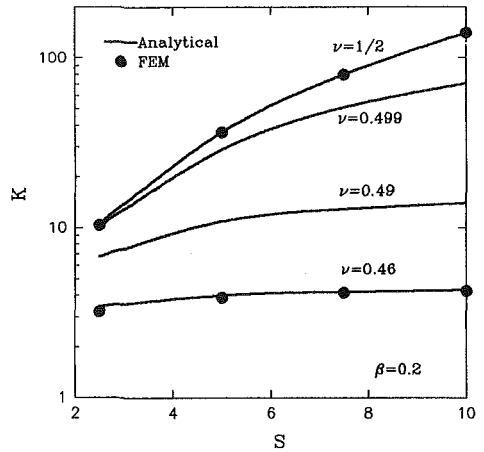


Fig. 9 Illustration of the relationship between stiffness K and shape factor S and Poisson's ratio ν for $\beta = 0.2$

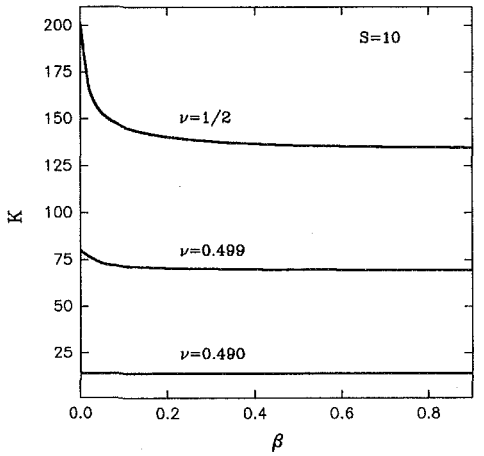


Fig. 10 Stiffness K versus radius β and Poisson's ratio ν for $S = 10$

6 Conclusions

The problem of the compression of a bonded thin annular disk has been solved approximately. It is based on the observation that for a thin disk, the boundary layers are confined to very narrow regions near the lateral surfaces. The so-called Perturbation-Ritz Method is proposed, in which the outer expansion is chosen as the trial functions, and the unknown constants

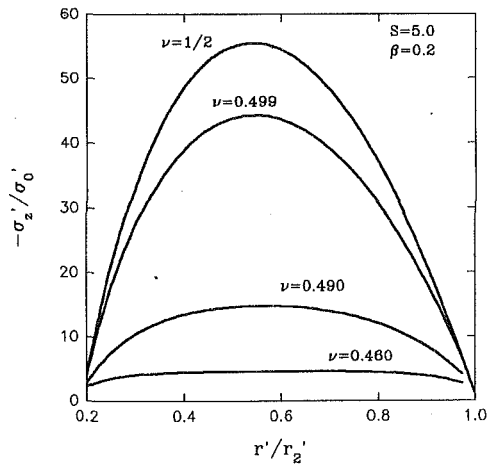


Fig. 11 Effect of Poisson's ratio ν on distribution of stress σ_z' at the plane end for $S = 10$ and $\beta = .2$. The normalization constant is $\sigma_0' = E'e$.

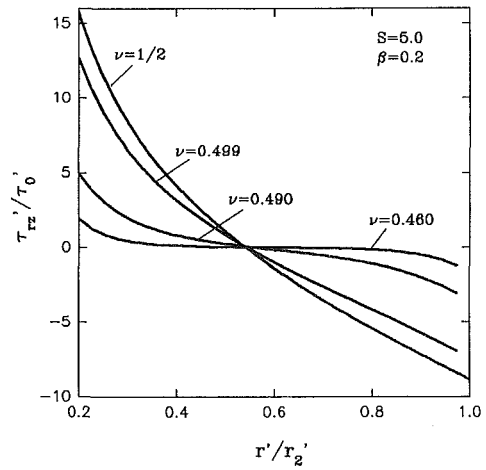


Fig. 12 Effect of the Poisson's ratio ν on distribution of stress τ_{rz}' at the plane end for $S = 10$ and $\beta = .2$. The normalization constant is $\tau_0' = \sigma_0' = E'e$.

or the coefficients are determined by the Ritz Method of minimizing the potential energy. The coefficients so obtained lead to much more accurate results than using Saint-Venant's Principle. The comparison of these analytical results with the finite element calculation results suggests that for large shape factors, say $S > 2$, the present solutions are sufficiently accurate for most engineering applications.

Poisson's ratio has a dramatic effect on the stiffness. Even slight compressibility causes significant reduction of stiffness from the incompressible case. There is a rapid change in stiffness near $\beta = 0$ in the incompressible case. However, as the material becomes more compressible, the change becomes more gradual, and the stiffness becomes almost independent of the radius ratio β. Thus, in the design of thin rubber disks in compression, the assumption of incompressibility is not appropriate, and it should be taken into account.

References

Benthem, J. P., and Minderhoud, P., 1972, "The Problem of the Solid Cylinder Compressed Between Rough Rigid Stamps," *Int. J. of Solids and Structures,* Vol. 8, pp. 1027–1042.

Chalhoub, M. S., and Kelly, J. M., 1991, "Analysis of Infinite-Strip-Shaped Base Isolator with Elastomer Bulk Compression," *ASCE J. of Engng. Mech.,* Vol. 117, pp. 1791–1805.

Fung, Y. C., 1965, *Foundations of Solid Mechanics,* Prentice-Hall, Englewood, Cliffs, NJ.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Geer, J. F., and Andersen, C. M., 1989, "A Hybrid Perturbation Galerkin Technique with Applications to Slender Body Theory," *SIAM J. Appl. Math.*, Vol. 49, pp. 344–361.

Geer, J. F., and Andersen, C. M., 1990, "A Hybrid Perturbation-Galerkin Technique that Combines Multiple Expansions," *SIAM J. Appl. Math.*, Vol. 50, pp. 1474–1495.

Gent, A. N., and Lindley, P. B., 1959, "The Compression of Bonded Rubber Blocks," *Proc. Inst. Mech. Engr., London,* Vol. 173, pp. 111–117.

Gent, A. N., et al., 1974, "Interfacial Stresses for Bonded Rubber Blocks in Compression and Shear," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 885–859.

Gregory, R. D., and Wan, F. Y. M., 1984, "Decaying States of Plane Strain in a Semi-Finite Strip and Boundary Conditions for Plate Theory," *J. of Elasticity,* Vol. 14, pp. 27–64.

Ling, Y., Engel, P. A., and Brodsky, W. L., 1992, "Nonlinear Load-Deflection Relations of Bonded Elastomeric Circular Disks and Strips," *Rubber Chemistry and Technology,* Vol. 65, pp. 917–931.

Ling, Y., Engel, P. A., and Geer, J. F., 1994, "The End Problem of Incompressible Elastic Cylinder," ASME JOURNAL OF APPLIED MECHANICS, Vol. 61, pp. 30–37.

Ling, Y., Engel, P. A., and Brodsky, W. L., 1995, "The Compression of Bonded Annular Rubber Blocks," *ASCE Journal of Engineering Mechanics,* Vol. 121, pp. 661–666.

Reiss, E. L., 1962, "Symmetric Bending of Thick Circular Plates," *SIAM J. Appl. Math.*, Vol. 10, pp. 596–609.

Robert, M., and Keer, L. M., 1987, "An Elastic Circular Cylinder with Displacement Prescribed at the End—Axially Symmetrical Case," *Q. J. Mech. Appl. Math.*, Vol. 40, pp. 339–363.

Swan, D. A., 1968, "The Semi-Infinite Cylinders with Prescribed End Displacement," *SIAM J. of Appl. Math.*, Vol. 16, pp. 860–881.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Intergranular Crack Nucleation in Bicrystalline Materials Under Fatigue

H. M. Shodja
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208
Assoc. Mem. ASME

Y. Hirose
Department of Material Science
and Engineering,
Kanazawa University,
Kanazawa, Japan

T. Mura
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208
Fellow ASME

During cyclic deformation of polycrystalline materials, as substantiated by many experimental observations, due to existence of high stress concentration at the interfaces the preferential site of crack nucleation is intercrystalline. Accordingly, it is assumed that the highly localized cyclic deformation persistent slip band (PSB) occurs along the grain boundary (GB) which results in intergranular crack initiation. In the present work the irreversible accumulation of dislocations are used to characterize PSB by means of double pile-up which are composed of vacancy and interstitial dipoles. We shall give the mechanism and a quantitative remedy of ratcheting of plastic deformation peculiar to fatigue deformation. In a manner conceptually analogous to Griffith theory (1921), the critical number of cycles to failure and hence the S-N curves for crack initiation is obtained by considering the free energy of the system. The Gibbs free energy change ΔG increases with the fatigue cycle number due to cyclic increment of elastic strain energy which in turn stems from cyclic pile-up of dislocations along the slip planes. The Gibbs free energy change attains its maximum value at a critical cycle number beyond which the state of dislocation dipole accumulation becomes energetically unstable. In our theory we postulate that this critical state is the onset of crack initiation. We shall give a quantitative value for the fatigue limit and analyze the dependence of the S-N curve on several important physical parameters such as grain size; surface energy; yield strength; width of the PSB; and the ratio of the shear modulus of the bicrystalline material.

1 Introduction

Deformation of bicrystalline materials is greatly affected by the presence of grain boundaries (GBs). To date, many experimental observations have confirmed that surfaces and interfaces such as GBs and twin boundaries are common sites for crack initiation. Westwood (1961) and Johnston et al. (1962), in their experimental studies reported the role of the GB in Magnesium Oxide bicrystals under compression. Johnston et al. studied the behavior of bicrystalls with wide range of misorientations in some details. In this paper, we shall consider, as an element of a bicrystalline material two perfectly bonded crystals with compatible GB that are free of any microscopic inhomogeneities such as impurities and voids. Figure 1 depicts two situations where crystals A and B are the constituents of bicrystalline materials. In Fig. 1(*b*), crystal A is totally surrounded by crystal B.

Forsyth (1953) was the first to predict and observe extrusion phenomena in the fatigue of aluminium-copper alloy. Later, as pointed out in the review articles of Grosskreutz (1971), Laird and Duquette (1972), and the references therein, the extrusion phenomena was also observed in a wide variety of materials including most of the f.c.c, b.c.c, and some h.c.p metals. The mechanism of crack nucleation, particularly under low strain amplitude fatigue, is primarily due to extrusion and intrusion from striations in grains adjacent to the free surface. Subsequently, this mechanism roughens the free surface, which are stress raisers on the surface and hence lead to crack initiation. Among other types of crack nucleation we should mention the Zener-Stroh crack, which occurs in coalescence of edge dislocations in a pile-up against an obstacle such as a GB. However, the latter is not commonly observed during fatigue experimentation. Fatigue processes is usually accompanied by narrow bands of highly localized cyclic strain, namely persistant slip bands (PSBs). In contrast to the Zener-Stroh crack, numerous experimental observations reveal that, for most materials, formation of cracks at GBs is the consequence of pile-up and coalescence of edge dislocation on two closely spaced layers with a band width much smaller than the length of the layer. This is a manifestation of existence of high tensile stress over a small area between the slip planes. Note that under fatigue, the material within the band becomes soft compared to the matrix and subsequent cyclic loading causes an accumulation of damage or equivalent ratcheting of plastic deformation inside the PSBs.

The preferred site of crack nucleation in bicrystalline materials, particularly under high strain amplitude fatigue, is intergranular or transgranular crack initiation at the GB. This is due to the high resolved shear stress concentration which exists there. At elevated temperatures, GBs serve as a crack nucleus, regardless of strain amplitude. At high plastic strain amplitude, $\Delta\epsilon_{pl}$, and high cyclic plastic strain rate, $\dot{\epsilon}_{pl}$, Mughrabi et al. (1981) attributed the nonoccurance of fatigue cracks in slip bands and the preferential intercrystalline crack initiation in polycrystalline α-iron to incompatible shape changes resulting from asymmetric slip in neighboring grains at the surface. On the other hand, under the same $\Delta\epsilon_{pl}$ but low $\dot{\epsilon}_{pl}$, the slip bands become more prominent and the mode of crack nucleation is predominantly transgranular (Mughrabi, 1975; Mughrabi et al., 1981). Nevertheless, in our theoretical model we shall assume that, the formation of PSB along the GB eventually leads to intercrystalline crack initiation. The dislocation dipole model has been used to simulate PSB and nucieation of microvoids at the interface between a thin film and a substrate (Mura, 1994; Qin et al. 1991). We shall use vacancy and interstitial dipoles to model PSBs. As illustrated in Figs. 2(*a*), (*b*), and (*c*) we predict

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Dec. 9, 1994; final revision, Aug. 1, 1995. Associate Technical Editor: D. M. Barnett.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

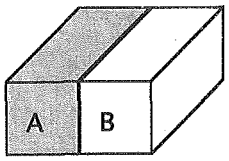


Fig. 1(*a*)

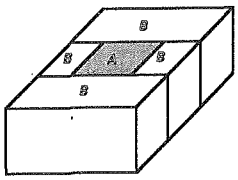


Fig. 1(*b*)

Fig. 1 Two examples of bicrystalline materials

three distinct possibilities for the formation of PSB with respect to the interface, with points O and O' indicating the internal dislocation sources, such as the Frank-Reed type source. Figure 3 shows the applied cyclic loading with τ_{max} and τ_{min} as the maximum and minimum shear stresses, respectively, with loading amplitude, $\Delta\tau = \tau_{max} - \tau_{min}$. We assume that, before loading, the crystals are free of any dislocations. When first loading to point 1, as shown in Fig. 3, the source O is activated emitting dislocation pairs of positive and negative signs along layer I. Similarly upon reversal of loading to point 2, the internal source O' emits dislocation pairs of both signs along layer II. The presence of some strong obstacles on either side of the internal sources in the path of layers I and II allow for the double pile-up of dislocations against the obstacles along these layers. Continued cyclic loading increases the dislocation density in the pile-up. Consequently, high density of vacancy dipoles at one end and high density of interstitial dipoles at the other end is equivalent to the movement of the material from the high density region of vaccancy dipoles across OO' toward the high density region of interstitial dipoles. Eventually, this process results in microvoid nucleation in the region with a high density of vacancy dipoles, and hillock formation in the region of high interstitial dipoles.

The criteria used for obtaining the critical cycle number for fatigue life and subsequently obtaining the S-N (S: applied stress amplitude, N: number of cycles to crack initiation) curve is adopted from Mura and Nakasone (1990) and is similar to Griffith theory (1921). The energy of a Griffith crack under a fixed applied stress continues to increase with crack length up to the (unstable) equilibrium crack length at which the total free energy attains its maximum. Similarly for the fatigue problem, for a given stress amplitude $\Delta\tau$, the Gibbs free-energy change ΔG is expressed as a function of cycle number n (Fig. 4). This free energy change increases with the fatigue cycle number due to the increase in elastic strain energy and becomes maximum at a critical cycle number beyond which the initial state of dislocation dipole accumulation becomes energetically unstable. The energy of this state will be released by annihilation of dislocation dipoles of vacancy type followed by formation of intergranular fatigue crack. Using these concepts we shall analyze the effect of several important physical parameters such as grain size; surface energy; yield strength; width of the PSB; and the ratio of the shear modulus of the crystals on the fatigue life of the bicrystal and give the corresponding S-N curves. One of the important features of these S-N curves that we gain in applying our theory is the fatigue limit, which indicates the threshold stress amplitude, below which cracks cannot be nucleated regardless of the number of cycles.

2 Mathematical Formulation of Dislocation Dipole Model

2.1 Equilibrium of Peach-Koehler Force and the Governing Integral Equation. Consider two perfectly bonded semi-infinite domains to represent an element of a bicrystalline material with three distinct situations, where slip planes are parallel to the interface, as depicted in Fig. 2. In this paper we focus on the case shown in Fig. 2(a). The remaining cases can be treated in a similar fashion. The material under consideration is elastic-perfectly plastic, so that the frictional stress τ_f, which is the stress necessary for moving a dislocation inside the crystal lattice is constant. At stage 1 of loading cycle (Fig. 3) the maximum applied shear stress τ_{max} is greater than τ_f and hence results in dislocation distribution on layer I. At the onset of

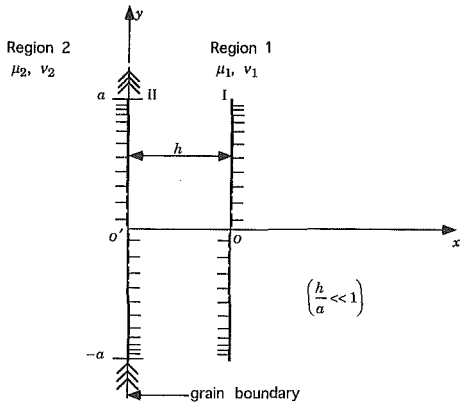


Fig. 2(*a*)

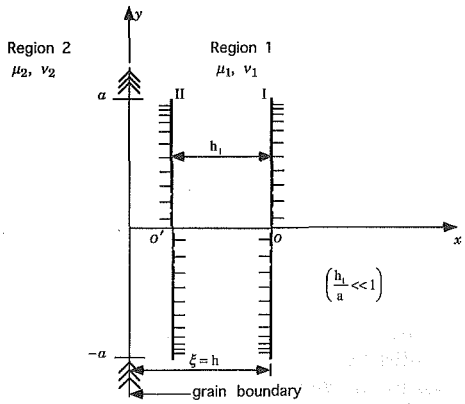


Fig. 2(*b*)

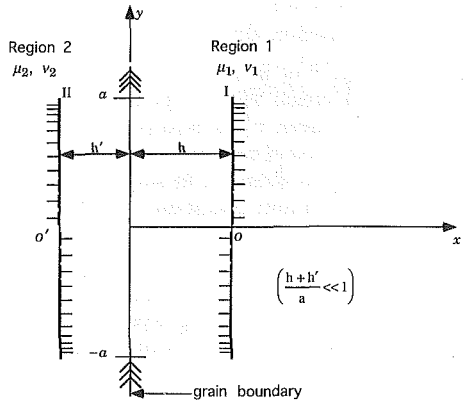


Fig. 2(*c*)

Fig. 2 Dislocation dipole model of the PSB at the GB, slip plane *II* is (*a*) at the GB (*b*) inside region 1 and (*c*) inside region 2

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

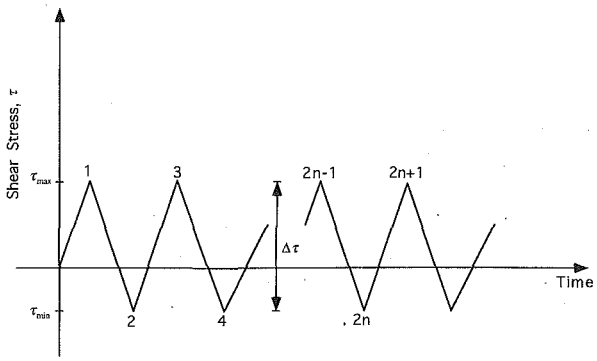


Fig. 3 Cyclic shear stress pattern acting on the slip planes versus time

plastic flow, the equilibrium of Peach-Koehler force on the piled-up dislocations along the slip plane I yields

$$\tau_1^D + \tau_{\max} = \tau_f, \tag{1}$$

where τ_1^D is the back stress (dislocation stress) due to the dislocation distribution on layer I at stage 1 of the loading cycle. While the dislocation stress τ_1^D is positive on layer I, it contributes a negative back stress, but nearly equal magnitude on layer II, which is located very close to layer I. The magnitude of the back stress τ_1^D decreases with distance from layer I. Therefore, in view of the above and the assumption that dislocation motion is irreversible, it is reasonable that, upon unloading, the reverse slip occurs very close to layer I. On the other hand, in order to avoid annihilation of dislocations, the minimum distance between layers I and II must be of dislocation cell size. Ad hoc the distance between the two layers is negligible as compared to the length of the pile-up.

Load reversal to point 2 of the loading cycle, with the corresponding minimum applied shear stress $\tau_{\min}$, gives rise to the dislocation stress τ_2^D due to new generation of dislocation distribution along layer II. The equilibrium of Peach-Koehler force on the piled-up dislocation along the newly formed slip plane II is given by

$$\tau_2^D + \tau_1^D + \tau_{\min} = -\tau_f. \tag{2}$$

As it appears from Eq. (2), the reverse plastic flow on layer II is indeed assisted by the back stress τ_1^D. In this manner, ratcheting of plastic deformation is incorporated into our theoretical model. The mechanism of the build up of large local plastic strains with cyclic loading within two closely located thin slices was described by Lin and Ito (1969); however, their study did not consider dislocations, which play a fundamental role.

Combining Eqs. (1) and (2) and utilizing $\Delta\tau = \tau_{\max} - \tau_{\min}$,

$$\tau_2^D = \Delta\tau - 2\tau_f. \tag{3}$$

It follows that at stage 2 of loading cycle, $\Delta\tau$ must be greater than $2\tau_f$ in order to create new dislocations along layer II. Noting that the back stress τ_2^D is positive on layer I, we see how this aids the formation of new dislocation distribution along the layer with dislocation stress τ_3^D at stage 3 of loading cycle and results in the equilibrium condition

$$\tau_3^D + \tau_{\max} + \tau_2^D = \tau_f. \tag{4}$$

Reversal of loading reactivates the dislocation source O', emitting additional dislocation pairs along layer II. The equilibrium condition along layer II and at stage 4 of the loading cycle is

$$\tau_4^D + \tau_{\min} + \tau_3^D = -\tau_f, \tag{5}$$

where τ_4^D is the back stress due to the dislocation distribution on layer II. At this point the trend of acceleration of the damage with each loading cycle is clearly seen from Eqs. (1), (2), (4), and (5), in which the current applied back stress is the accumulation of all the previous ones. With the aid of Eqs. (1) and (3), Eq. (4) may be rewritten as

$$\tau_3^D - \tau_1^D = -(\Delta\tau - 2\tau_f). \tag{6}$$

Similarly, Eqs. (4) and (5) yield

$$\tau_4^D - \tau_2^D = \Delta\tau - 2\tau_f. \tag{7}$$

Following the above pattern, after n loading cycles, the equilibrium of the Peach-Koehler force on the piled-up dislocations yield

$$\tau_{2n+1}^D - \tau_{2n-1}^D = -(\Delta\tau - 2\tau_f), \quad \text{on layer } I, \quad x = h,$$

$$\tau_{2n}^D - \tau_{2n-2}^D = \Delta\tau - 2\tau_f, \quad \text{on layer } II, \quad x = 0.$$

However, for a large cycle number n of fatigue loading and unloading, $\tau_1^D/n \to 0$, so that

$$\tau_{2n+1}^D \approx -n(\Delta\tau - 2\tau_f), \quad \text{on layer } I, \quad x = h, \tag{8}$$

$$\tau_{2n}^D = n(\Delta\tau - 2\tau_f), \quad \text{on layer } II, \quad x = 0. \tag{9}$$

Equations (8) and (9) yield at once

$$\tau_{2n+1}^D(y) \approx -\tau_{2n}^D(y).$$

That is, after a large cycle number of fatigue, the dislocation stresses, and consequently the dislocation distribution densities on the two layers, are approximately equal, but are of opposite signs. Hence

$$D_{2n+1}(y) \approx -D_{2n}(y),$$

where the dislocation density distribution $D_{2n+1}(y)$ on layer I is obtained from the following integral equation:

$$\int_{-a}^{a} D_{2n+1}(\eta)\tau_{xy}(h, y; h, \eta)d\eta = -n(\Delta\tau - 2\tau_f), \tag{10}$$

where $\tau_{xy}(h, y; h, \eta)$ is the shear stress field at (h, y) due to an edge dislocation at (h, η). τ_{xy} is given in the following subsection.

2.2 Stress Field of an Edge Dislocation Near the Grain Boundary. The problems of the interaction of dislocations and boundaries, and edge dislocations in inhomogeneous media were first studied by Head (1953a, b). Consider two perfectly bonded elastic half-spaces with shear moduli μ_1, μ_2 and Pois-

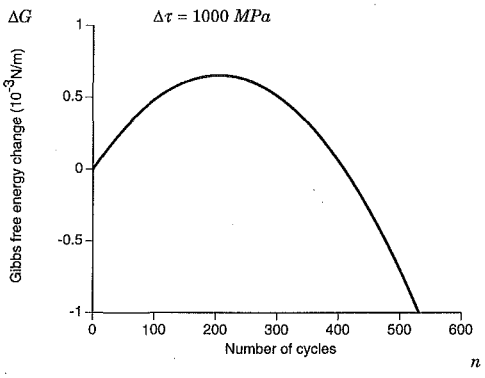


Fig. 4 Typical plot of Gibbs free-energy change versus cycle number. Critical cycle number occurs when ΔG attains its maximum.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

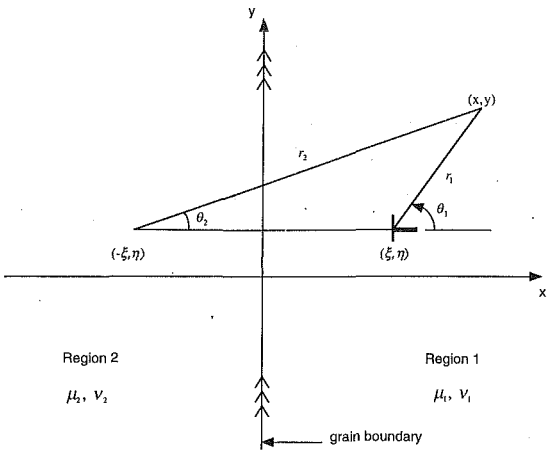


Fig. 5 An edge dislocation near an interface lacated at (ξ, η) in region 1

son's ratios ν_1, ν_2. As depicted in Fig. 5, a single edge dislocation is located at (ξ, η) in region 1 of the two joined half-spaces with its Burgers vector parallel to the interface. The coordinates and the geometrical location of the edge dislocation, with respect to its image at $(-\xi, \eta)$ and the field point (x, y) are also shown in Fig. 5. When an edge dislocation is located at $(\xi, 0)$ in region 1 with its Burgers vector parallel to the interface, the Airy stress functions for the two regions are given by (Dundurs, 1969)

$$\chi^{(1)} = \frac{2\mu_1 b_y}{\pi(\kappa_1 + 1)(1 - \beta^2)} \Big\{ (1 - \beta^2)(r_1 \log r_1) \cos\theta_1 + (\alpha + \beta^2)(r_2 \log r_2) \cos\theta_2 - (1 + \alpha)\beta r_2 \theta_2 \sin\theta_2 - (\alpha - \beta)(1 - \beta) \times \xi\left(2 \log r_2 - \cos 2\theta_2 + 2\xi \frac{\cos\theta_2}{r_2}\right)\Big\}, \quad (11)$$

$$\chi^{(2)} = \frac{2\mu_1 b_y (1 + \alpha)}{\pi(\kappa_1 + 1)(1 - \beta^2)} \{(r_1 \log r_1) \cos\theta_1 + \beta(r_1\theta_1 \sin\theta_1 + 2\xi \log r_1)\}, \quad (12)$$

where b_y is the Burgers vector, α and β are Dundurs constants

$$\alpha = \frac{\Gamma(\kappa_1 + 1) - (\kappa_2 + 1)}{\Gamma(\kappa_1 + 1) + \kappa_2 + 1},$$

$$\beta = \frac{\Gamma(\kappa_1 - 1) - (\kappa_2 - 1)}{\Gamma(\kappa_1 + 1) + \kappa_2 + 1},$$

with $\Gamma = \mu_2/\mu_1$, $\kappa_1 = 3 - 4\nu_1$ and $\kappa_2 = 3 - 4\nu_2$. Employing the Airy stress function given by Eq. (11) and noting that $\tau_{xy}(x, y; \xi, \eta) = \tau_{xy}(x, y - \eta; \xi, 0)$ it follows that

$$\tau_{xy}(x, y; \xi, \eta) = \frac{2\mu_1 b_y}{\pi(\kappa_1 + 1)(1 - \beta^2)} \times \Big\{ -(1 - \beta^2)\frac{\sin\theta_1 \cos 2\theta_1}{r_1} - (\alpha + \beta^2)\frac{\sin\theta_2 \cos 2\theta_2}{r_2} + 2\beta(1 + \alpha)\frac{\sin\theta_2 \cos^2\theta_2}{r_2} - \xi(\alpha - \beta)(1 - \beta)\left(-\frac{2\sin 2\theta_2}{r_2^2} - \frac{2\sin 4\theta_2}{r_2^2} + \frac{4\xi \sin\theta_2 (4\cos^2\theta_2 - 1)}{r_2^3}\right)\Big\}. \quad (13)$$

In Eq. (13), r_1, r_2, θ_1, and θ_2 are redefined according to Fig. 5. Combination of Eqs. (10) and (13) yield

$$\frac{1}{\pi} {}^{*}\!\int_{-a}^{a} \frac{D_{2n+1}(\eta)}{\eta - y}\, d\eta - \frac{1}{\pi}\int_{-a}^{a} k(y, \eta) D_{2n+1}(\eta) d\eta = A, \quad (14)$$

with

$$A = \frac{n(\kappa_1 + 1)(\Delta\tau - 2\tau_f)}{2\mu_1 b_y}, \quad (15)$$

$$k(y, \eta) = -\frac{\alpha + \beta^2}{8h(1 - \beta^2)} \sin 4\theta_2 + \frac{\alpha + \beta}{4h(1 - \beta)} \cos^2\theta_2 \sin 2\theta_2 + \frac{\alpha - \beta}{4h(1 + \beta)} \sin 4\theta_2 \cos^2\theta_2, \quad (16)$$

where $-a \le (y, \eta) \le a$. The first term in Eq. (14) has a Cauchy-type singularity $1/(y - \eta)$ and the asterisk (*) on the upper left corner of the integral sign means that the integral exists in the sense of a Cauchy principal value only. Equation (14) is a Fredholm integral equation of the first kind, with a generalized Cauchy type kernel $1/(y - \eta)$ and a bounded Fredholm kernel $k(y, \eta)$ and is solved when subjected to the condition

$$\int_{-a}^{a} D_{2n+1}(\eta) d\eta = 0. \quad (17)$$

This additional condition assures the uniqueness of the solution, and states that the net Burgers vector of the dislocation density distribution on the slip planes equals to zero. Due to the singular behavior of the integral Eq. (14) and its complicated Fredholm kernel, an analytical solution of Eq. (14), when subjected to the constraint Eq. (17), is difficult to obtain. To circumvent this difficulty, and to preserve the nature of the power singularity of the dislocation density distribution, we employ a special numerical scheme which we describe in the following section.

3 Dislocation Density Distribution

The type of singular integral equation which we described in the previous section requires the use of an appropriate quadrature scheme. In particular, the Gauss-Jacobi integration technique, which has been extended by Erdogan and Gupta (1972), provides a powerful scheme suitable for the singular integral equations of the type developed herein. To employ this scheme, we shall make use of a proper transformations of variables y and η, such that the integration limits $(-a, a)$ of Eqs. (14) and (17) are transformed to the normalized interval $(-1, 1)$. Subsequently, discritization of Eqs. (14) and (17) leads to the following system of linear algebraic equations in dislocation density distribution

$$\sum_{i=1}^{m} \frac{1}{m} \mathcal{D}(t_i')\left[\frac{1}{t_i' - t_j} + \pi \mathcal{K}(t_i', t_j)\right] = A, \quad j = 1, 2, \ldots, m - 1, \quad (18)$$

$$\sum_{i=1}^{m} \mathcal{D}(t_i') = 0. \quad (19)$$

The right-hand side A and the function $\mathcal{K}$ are Hölder continuous in $-1 \le (t', t) \le 1$

$$\mathcal{K}(t', t) = -\frac{a}{\pi} k(t', t),$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

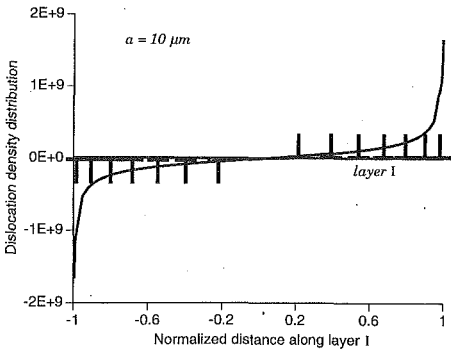


Fig. 6 Dislocation density distribution along the slip plane *I*

with θ_2 redefined as

$$\theta_2 = \tan^{-1}\frac{a(t - t')}{2h}.$$

$\mathcal{D}(t')$ is a bounded function and is related to $D_{2n+1}(t')$ through

$$\mathcal{D}(t') = \frac{D_{2n+1}(t')}{w(t')},$$

where the weight function $w(t')$ has the same type of singularity at the end points of the interval $-1 \le t' \le 1$ as $D_{2n+1}(t')$ and is defined by

$$w(t') = \frac{1}{\sqrt{1 - t'^2}}.$$

Furthermore,

$$t_i' = \cos\frac{(2i - 1)\pi}{2m}, \quad i = 1, 2, \ldots, m,$$

$$t_j = \cos\frac{j\pi}{m}, \quad j = 1, 2, \ldots, m - 1.$$

Figure 6 shows the graph of a typical dislocation density distribution along normalized distance on layer I, which was obtained for a pile-up with half-length equal to 10 μm. As shown in Fig. 6, the dislocation density distribution is equal to zero at the center of the slip plane and varies gradually away from it; near the tips of the slip plane, however, it varies rather abruptly and eventually becomes infinite right at the tips.

4 Gibbs Free-Energy Change and the S-N curves

The increase in Gibbs free-energy change ΔG with the fatigue cycle number is due to cyclic increments of elastic strain energy, which in turn stems from cyclic pile-up of dislocations along the slip planes. The Gibbs free-energy change, from a state of dislocation dipole accumulation (Fig. 2(a)) to the state of crack initiation of size c (Fig. 7), is

$$\Delta G = -W_1 - W_2 + 2c\gamma_s, \tag{20}$$

where γ_s is the surface energy, W_1 is the elastic strain energy release associated with the vacancy type dislocation dipoles, and W_2 is the mechanical energy release due to the opening up of a crack embryo of size c. The stored energies W_1 and W_2 will be released when a crack is nucleated, while $2c\gamma_s$ is the energy necessary to create the two faces of the microvoid of size c

$$c = b_y N(a) = b_y \int_0^a D_{2n+1}(\eta)d\eta,$$

where $N(a)$ is the total number of positive dislocations on layer I, equivalently in terms of the normalized variable introduced in previous section

$$c = ab_y \int_0^1 D_{2n+1}(t')dt'. \tag{21}$$

Since the mechanical energy release is much less than the elastic strain energy release, we shall discount W_2. The expression that relates elastic strain energy release to the self and interaction energies is given in the following subsection.

4.1 Elastic Strain Energy; Relation to Interaction and Self Energies. The stored elastic strain energy due to dislocation dipole accumulation may be written as

$$W_1 = \frac{1}{2}\int_\Omega \sigma_{ij}^D u_{i,j}^D dV, \tag{22}$$

where σ_{ij}^D and $u_{i,j}^D$ are the dislocation stress and elastic distortion, respectively. dV is an elemental volume within the entire domain Ω. When we employing the divergence theorem and note $\sigma_{ij,j} = 0$ in Ω, Eq. (22) becomes

$$W_1 = \frac{1}{2}\int_{\partial\Omega} \sigma_{ij}^D u_i^D n_j dS, \tag{23}$$

where dS is a surface element on the boundary $\partial\Omega$ with outer unit normal n_j. Since the slip takes place only along layers I and II, with nonzero tractions, Eq. (23) yields

$$W_1 = -\int_{-a}^{a} \tau_{2n+1}^D(h, y)[u_2^D(y)]_I dy - \int_{-a}^{a} \tau_{2n+1}^D(0, y)[u_2^D(y)]_{II} dy, \tag{24}$$

where $\tau_{2n+1}^D(h, y)$ and $\tau_{2n+1}^D(0, y)$ are the dislocation stresses caused by dislocation distribution along layer I. However, $\tau_{2n+1}^D(h, y)$ is acting along layer I whereas $\tau_{2n+1}^D(0, y)$ is acting on layer II. The slip distance $[u_2^D(y)]_I$ at a point y on layer I is given by

$$[u_2^D(y)]_I = (u_2^{D+}(y))_I - (u_2^{D-}(y))_I = b_y\{N(a) - N(|y|)\} = b_y\int_{|y|}^{a} D_{2n+1}(\eta)d\eta,$$

for large enough cycle number n of fatigue loading and unloading the slip distance along layer II, $[u_2^D(y)]_{II} \approx$

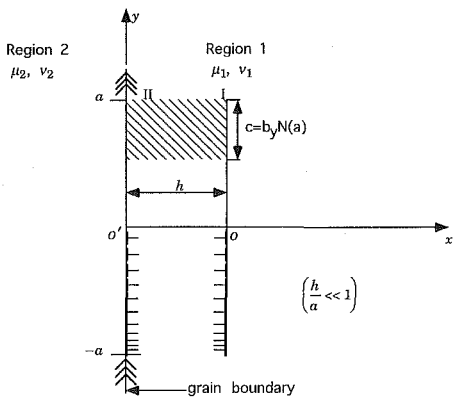


Fig. 7 Annihilation of dislocation dipoles of vacancy type followed by formation of intergranular crack embryo of size *c*

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

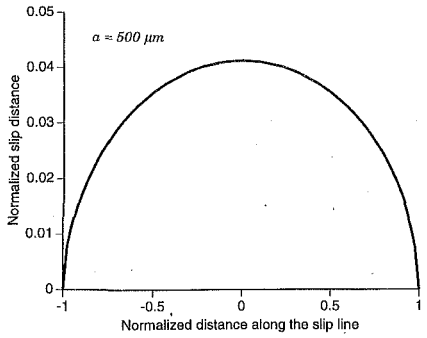


Fig. 8 Slip distance along the slip plane *I*

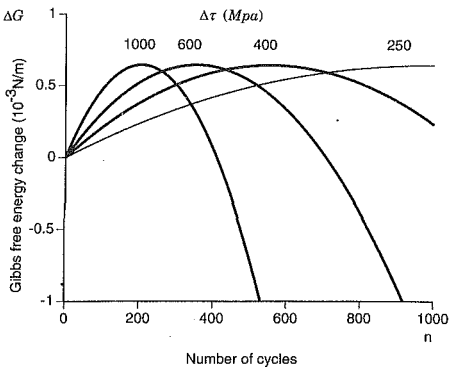


Fig. 9 Computed Gibbs free-energy change versus number of cycles for different applied shear stress amplitudes

$-[u_2^D(y)]_I$. Using the transformation of variables which we introduced in Section 3

$$[u_2^D(t)]_I = ab_y \int_{|t|}^{1} D_{2n+1}(t')dt'. \tag{25}$$

Recall that dislocation density distribution is antisymmetric (Fig. 6), whereas in Fig. 8, the slip distance versus normalized distance along the slip plane is symmetric.

The first and second terms on the right-hand side of Eq. (24) are the self and the interaction energies, respectively. We denote the self energy by W_{self} and those of interaction energy by W_{int}, thus we may rewrite Eq. (24) as

$$W_1 = 2(W_{\text{self}} + W_{\text{int}}). \tag{26}$$

With the aid of the transformation of the variables introduced in Section 3, the self and interaction energies for layer *I* becomes

$$W_{\text{self}} = -\tfrac{1}{2}\, a \int_{-1}^{1} \tau_{2n+1}^D(h, t)[u_2^D(t)]_I dt, \tag{27}$$

$$W_{\text{int}} = -\tfrac{1}{2}\, a \int_{-1}^{1} \tau_{2n+1}^D(0, t)[u_2^D(t)]_{II} dt. \tag{28}$$

Since $[u_2^D]_{II} \approx -[u_2^D]_I$ and $\tau_{2n}^D \approx -\tau_{2n+1}^D$, it turns out that W_{self} and W_{int} for layers *I* and *II* are approximately equal, thus the factor 2 in Eq. (26) accounts for both layers.

4.2 Crack Nucleation. From Eqs. (18) and (19) we have

$$\mathbf{D}_{2n+1}(\mathbf{t}') = \mathcal{C}n(\Delta\tau - 2\tau_f)\boldsymbol{\Psi}^{-1}(\mathbf{t}', \mathbf{t})\boldsymbol{\Upsilon}, \tag{29}$$

where $\boldsymbol{\Psi}_{m\times m}$ is the coefficient matrix, $\boldsymbol{\Upsilon}_{m\times 1} = \{1, 1, \ldots, 1, 0\}^T$ and the constant $\mathcal{C}$ contains some materials parameters. It follows that

$$\Delta G = c_1 n^2(\Delta\tau - 2\tau_f)^2 + c_2 n(\Delta\tau - 2\tau_f)\gamma_s,$$

where the constants c_1 and c_2 contain information pertinent to material properties and geometry. The Gibbs free energy change ΔG attains its maximum at a critical cycle number n_{cr}, beyond which the initial state of dislocation dipole accumulation becomes energetically unstable

$$n_{cr} = \frac{c_2\gamma_s}{2c_1(\Delta\tau - 2\tau_f)}, \tag{30}$$

hence the *S-N* curves for fatigue crack initiation can be obtained. In plotting the *S-N* curves, $n_{cr} \equiv n_i$ determines the number of cycles to fatigue crack initiation. We assume that the embryonic crack of length c, which is created by the loading cycle number n_i becomes the unstable Griffith crack, which in turn extends to infinitely long crack, resulting in the fatigue failure. The plotting of Gibbs free-energy change ΔG versus the number of cycles for different applied shear stress amplitudes are shown in Fig. 9; we see that the critical cycle number increases with decreasing applied shear stress amplitude. The fact that the maximum Gibbs free energy change becomes less pronounced at lower applied shear stresses suggests a wider variation of fatigue life expectancy. It is interesting to note that the surface energy varies linearly with the critical cycle number, and has a square root dependence on the maximum Gibbs free-energy change.

5 Results and Discussion

According to the criteria set forth herein for obtaining the critical cycle number to fatigue life, we examine the effect of several important physical parameters on the *S-N* curves. As it appears in Eq. (30), our theory predicts that the threshold stress amplitude (fatigue limit) below which cracks cannot be nucleated is twice the frictional stress for the dislocation motion. Since the failure mechanism is the one by slip along the PSBs, materials with higher frictional stress naturally exhibit a higher yield strength, which results in higher fatigue life, in agreement with our findings. The phenomenon of dependence of the *S-N* curve on the frictional stress is illustrated in Fig. 10, where for a given $\Delta\tau$, we observe higher fatigue strength to crack initiation in the materials with higher τ_f. In Fig. 10 the difference in the fatigue limits as the applied shear stress approaches $2\tau_f$ is clearly demonstrated.

The effect of the surface energy on the *S-N* curve is shown in Fig. 11, where an increase in γ_s results in higher fatigue resistance. As expected, the higher the surface energy the higher the number of cycles to failure and consequently higher energy is required for creation of two faces of a crack embryo, which further substantiates our theoretical results.

One of the important parameters considered in this study is the length of the slip plane which is representative of a grain size. From Fig. 12, we observe that materials with larger grain

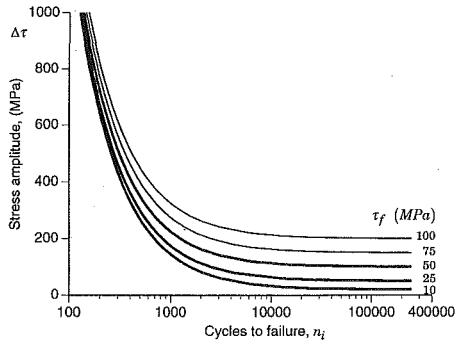


Fig. 10 Dependence of the *S-N* curve and the corresponding fatigue limit on the frictional stress

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

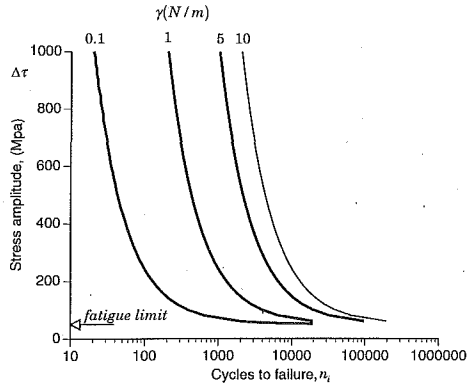


Fig. 11 Dependence of the *S-N* curve on the surface energy

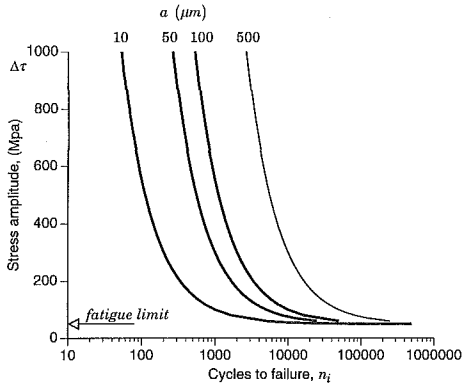


Fig. 12 Effect of the grain size on the fatigue life

size are stronger and have higher fatigue resistance. This is in agreement with experimental observations. Nakai et al. (1981) studied the effects of grain size on near-threshold fatigue crack propagation in low-carbon steel. The experimental results of Nakai et al., which is shown in Fig. 14, reveals that larger grain size materials correspond to higher threshold stress intensity range, and consequently have higher fatigue life. This phenomena was also confirmed by Gray et al. (1983). The microfractographic studies of fracture surfaces on the 7075 alloy in vacuum, which were performed by Petit (1983), indicate that the threshold fatigue crack growth rate is analagous to the fatigue crack initiation rate. This particular result of his experiment is shown in Fig. 15. Figure 13 depicts the solution of system of Eqs. (18) and (19) for the dislocation density distribution emitted from a single dislocation source for a fixed applied shear stress range

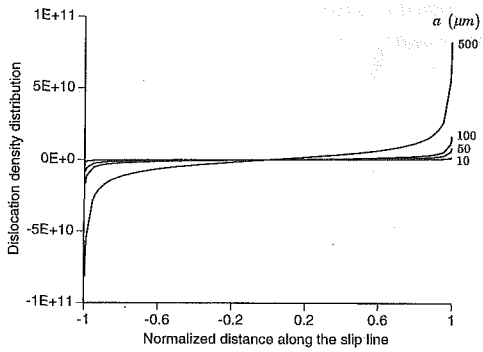


Fig. 13 Dislocation density distribution versus normalized distance along the slip plane. Dislocation pile-up shown for different half-length of slip planes correspond to the *S-N* curves shown in Fig. 12

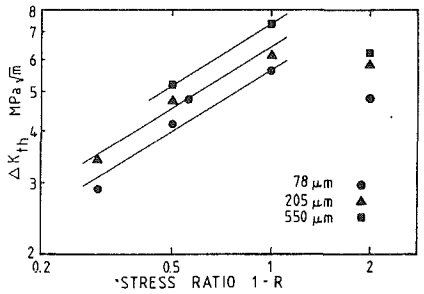


Fig. 14 Threshold stress intensity range, ΔK_{th} as a function of stress ratio, 1–R, and effect of grain size in a low-carbon steel (Nakai et al., 1981)

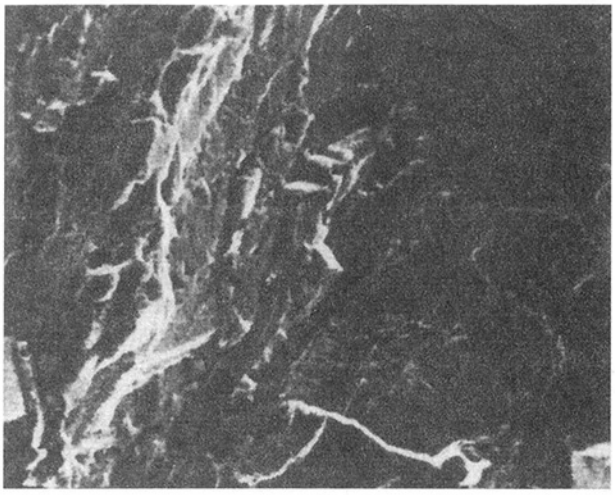


Fig. 15 Microfractographic view of fracture surface on the 7075 alloy in vacuum (Petit, 1983)

$\Delta\tau = 1$ Gpa and for slip planes of half length equal to 10, 50, 100, and 500 μm, which correspond to the *S-N* curves shown in Fig. 12.

Under cyclic deformation, mutual annihilations of dislocations play an important role during glide. Essmann and Mughrabi (1979) studied the critical height of the dislocation dipoles of vacancy type of 1.6 nm for copper below which the annihilation of edge dislocations occurs. In weak beam TEM study of fatigued copper single crystals, Antonopoulos et al. (1976) reported edge-dislocation dipole widths from 2 to 5 nm with a mean of 3.5 nm. As depicted in Fig. 16, we considered a range of PSB widths h equal to 0.5, 1, 2.5 and 5 nm in order to demonstrate the qualitative effect of PSB width on the *S-N*

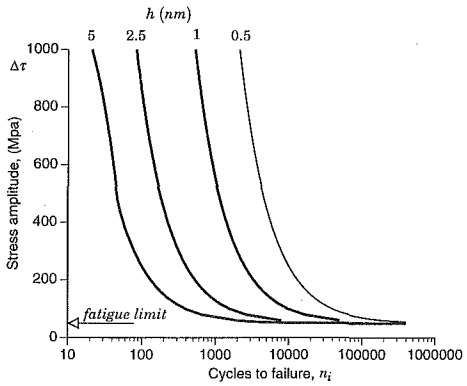


Fig. 16 Dependence of the *S-N* curve on the width of the PSB

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

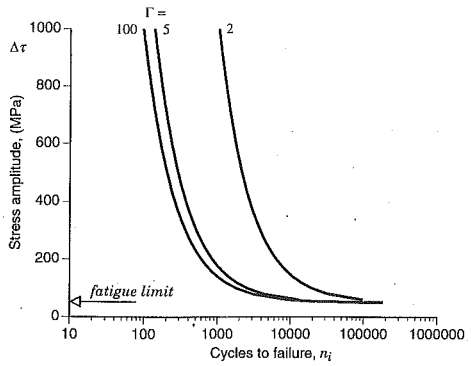


Fig. 17 Effect of ratio of the shear modulus of crystal 2 to those of crystal 1 ($\Gamma = \mu_2/\mu_1$) on the fatigue life of the bicrystalline material

curves. We infer that, among the coexisting dislocation dipoles of various widths, which are shown in Fig. 16, for the largest band width $h = 5$ nm it takes the least number of loading cycles to crack initiation.

The dependence of the *S-N* curves on the mismatch of the shear modulus is shown in Figs. 17; we conclude that larger mismatch of the shear modulus of grains results in lower fatigue resistance. In other words, we can say that the homogeneous materials have higher fatigue resistance to crack initiation. We close this section by noting that Figs. 11, 12, 16, and 17 were all obtained for the same value of friction stress $\tau_f = 25$ Mpa; therefore the same fatigue limit of 50 Mpa prevails for these cases.

6 Conclusions

Based on the existing studies, we proposed a model for dislocation dipole pile-up peculiar to intergranular fatigue crack initiation in bicrystalline materials. The phenomena of ratcheting of plastic deformation within two closely located thin slices or equivalently within the so-called PSBs, and acceleration of the damage with each loading cycle, were incorporated in our theoretical formulation through consideration of back stress.

We have shown that the dislocation density distribution along the slip plane is minimal away from the tip of the pile-up and becomes infinite right at the tip. The maximum Gibbs free-energy change corresponds to a critical cycle number, beyond which crack initiates at the tip of the pile-up of vacancy dipoles. We thoroughly analyzed the effect of the grain size, surface energy, yield strength, width of the PSB and the ratio of the shear modulus of the bicrystalline material on its fatigue life and calculated the corresponding *S-N* curves. Finally, based on our theory presented herein, we were able to obtain the fatigue limit, an important feature of the *S-N* curves.

Acknowledgments

This research was supported by AFOSR-F49620-92-J-0319.

References

Antonopoulos, J. G., Brown, L. M., and Winter, A. T., 1976, "Vacancy Dipoles in Fatigued Copper," *Philosophical Magazine,* Vol. 34, pp. 549–563.

Dundurs, J., 1969, "Elastic Interaction of Dislocations with Inhomogeneities," *Mathematical Theory of Dislocations,* T. Mura, ed., pp. 70–115.

Erdogan, F., and Gupta, G. D., 1972, "On the Numerical Solution of Singular Integral Equations," *Quarterly of Applied Mathematics,* Vol. 30, pp. 525–534.

Essmann, U., and Mughrabi, H., 1979, "Annihilation of Dislocations During Tensile and Cyclic Deformation and Limits of Dislocation Densities," *Philosophical Magazine,* Vol. 40, pp. 731–756.

Forsyth, P. J. E., 1953, "Exudation of Material from Slip Bands at the Surface of Fatigued Crystals of an Aluminum-Copper Alloy," *Nature,* Vol. 171, pp. 172–173.

Gray, III, G. T., Thompson, A. W., and Williams, J. C., 1983, "The Effect of Microstructure on Fatigue Crack Path and Crack Propagation Rate," *Fatigue Crack Growth Threshold Concepts,* D. Davidson and S. Suresh, eds., The Metallurgical Society of AIME, Oct. 3–5, pp. 131–143.

Griffith, A. A., 1921, "The Phenomena of Rupture and Flow in Solids," *Transactions of the Royal Society of London,* Vol. A221, pp. 163–179.

Grosskreutz, J. C., 1971, "The Mechanisms of Metal Fatigue," *Physica Status Solidi,* Vol. 47, pp. 359–396.

Head, A. K., 1953a, "The Interaction of Dislocations and Boundaries," *Philosophical Magazine,* Vol. 44, pp. 92–94.

Head, A. K., 1953b, "Edge Dislocations in Inhomogeneous Media," *The Proceedings of the Physical Society, London,* Vol. B66, pp. 793–801.

Johnston, T. L., Stokes, R. J., and Li, C. H., 1962, "Crack Nucleation in Magnesium Oxide Bi-crystals Under Compression," *Philosophical Magazine,* Vol. 7, pp. 23–34.

Laird, C., and Duquette, D. J., 1972, "Mechanisms of Fatigue Crack Nucleation," *Corrosion Fatigue,* O. F. Devereux, A. J. McEvily, and R. W. Staehle, eds., NACE 2, pp. 88–117.

Lin, T. H., and Ito, Y. M., 1969, "Mechanics of a Fatigue Crack Nucleation Mechanism," *Journal of the Mechanics and Physics of Solids,* Vol. 17, pp. 511–523.

Mughrabi, H., 1975, "Wechselverformung und Ermüdungsbruch von α-Eisenvielkristallen," *Zeitschrift fur Metallkunde,* Vol. 66, pp. 719–724.

Mughrabi, H., Herz, K., and Stark, X., 1981, "Cyclic Deformation and Fatigue Behaviour of α-Iron Mono- and Polycrystals," *International Journal of Fracture,* Vol. 17, pp. 193–220.

Mura, T., 1994, "A Theory of Fatigue Crack Initiation," *Materials Science and Engineering,* Vol. A176, pp. 61–70.

Mura, T., and Nakasone, Y., 1990, "A Theory of Fatigue Crack Initiation in Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp. 1–6.

Nakai, Y., Tanaka, K., and Nakanishi, T., 1981, "The Effects of Stress Ratio and Grain Size on Near-Threshold Fatigue Crack Propagation in Low-Carbon Steel," *Engineering Fracture Mechanics,* Vol. 15, No. 3–4, pp. 291–302.

Petit, J., 1983, "Some Aspects of Near-Threshold Crack Growth: Microstructural and Environmental Effects," *Fatigue Crack Growth Threshold Concepts,* D. Davidson and S. Suresh, eds., The Metallurgical Society of AIME, Oct. 3–5, pp. 3–24.

Qin, S., Fan, H., and Mura, T., 1991, "Microvoid Nucleation at the Interface Between a Thin Film and a Substrate in Fatigue," *Journal of Applied Physics,* Vol. 70, pp. 1405–1411.

Westwood, A. R. C., 1961, "On the Fracture Behaviour of Magnesium Oxide Bi-crystals," *Philosophical Magazine,* Vol. 6, pp. 195–200.

APPENDIX

In this appendix we shall give an alternative derivations of Eqs. (8) and (9). Consider the arbitrary stages $2j - 2$, $2j - 1$, $2j$ and $2j + 1$ of the loading cycles, then the corresponding equilibrium of Peach-Koehler force on the pile-up dislocations are

$$\tau^D_{2j-1} + \tau^D_{2j-2} + \tau_{\max} = \tau_f, \text{ on layer } I, \quad \text{stage } 2j - 1, \quad \text{(A1)}$$

$$\tau^D_{2j} + \tau^D_{2j-1} + \tau_{\min} = -\tau_f, \quad \text{on layer } II, \quad \text{stage } 2j, \quad \text{(A2)}$$

$$\tau^D_{2j+1} + \tau^D_{2j} + \tau_{\max} = \tau_f, \quad \text{on layer } I, \quad \text{stage } 2j + 1, \quad \text{(A3)}$$

from which we obtain

$$\tau^D_{2j} - \tau^D_{2j-2} = \Delta\tau - 2\tau_f, \quad \text{(A4)}$$

$$\tau^D_{2j+1} - \tau^D_{2j-1} = -(\Delta\tau - 2\tau_f). \quad \text{(A5)}$$

That is the net increment in the back stress between two consecutive even numbered stages (i.e., loading from $\tau_{\min}$ to $\tau_{\max}$ followed by unloading to $\tau_{\min}$) and those of odd numbered stages (i.e., unloading from $\tau_{\max}$ to $\tau_{\min}$ followed by loading to $\tau_{\max}$) are equal to $\Delta\tau - 2\tau_f$ and $-(\Delta\tau - 2\tau_f)$, respectively. Hence, the net increments are invariant of the cycle number. In view of the above, it takes n increments for the even numbered stages to reach to stage $2n$, and starting from stage 1 it takes n increments for the odd numbered stages to reach to stage $2n + 1$. After large enough cycle number τ^D_1 will be negligible and we obtain

$$\tau^D_{2n+1} \approx -n(\Delta\tau - 2\tau_f), \quad \text{on layer } I, \quad \text{stage } 2n + 1, \quad \text{(A6)}$$

$$\tau^D_{2n} = n(\Delta\tau - 2\tau_f), \quad \text{on layer } II, \quad \text{stage } 2n, \quad \text{(A7)}$$

which are Eqs. (8) and (9) given in Subsection 2.1.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Diffusive Crack Growth at a Bimaterial Interface

Tze-jer Chuang
Ceramics Division,
National Institute of Standards
and Technology,
Gaithersburg, MD 20899
Fellow ASME

June-Liang Chu
Graduate Student.

Sanboh Lee
Professor.

Department of Materials Science
and Engineering,
National Tsing Hua University,
Hsin-Chu, Taiwan 30043, R.O.C.

The high temperature microcrack growth behavior along a planar interface between two elastic dissimilar media is investigated with an aim at estimating service life of advanced ceramic composites under creep-rupture conditions. The crack is assumed to grow along the interface normal to a remote applied tensile stress via a coupled surface and grain-boundary diffusion under steady-state creep conditions. The crack-tip conditions were first derived from the asymmetric tip morphology developed by surface self-diffusion. The governing integro-differential equation containing the unknown tensile stress distribution along the interface ahead of the moving crack tip was derived and it was found that a new length parameter exists as a scaling factor for the interface for which the solution becomes identical to that of the single-phase media when plotted on the nondimensional physical plane. In contrast to the elastic stress solution which shows singularity at the tip and oscillatory character away from the tip, the creep stresses have a peak value away from the tip due to a wedging effect and interfacial sliding eliminates stress oscillation resulting in a decoupling between mode I and mode II stress fields. This stress solution ties the far-field loading parameter to the crack-tip conditions in terms of the unknown crack velocity to give a specific V-K *functional relationship. It was shown that a stress exponent of 12 in the conventional power-law crack growth emerges at higher applied stress levels. An analysis on energy balance shows that the energy release during crack growth amounts to the* J*-integral which derives mostly from work done by "wedging," not from strain energy loss. A constraint on interfacial diffusivities of the two species was found and its implications on possible microstructural developments were discussed.*

Introduction

The present paper is concerned with the crack growth behavior in an interface between two dissimilar phases subjected to long-term sustained loading conditions at elevated temperatures. In recent years, advanced ceramic monolithics and composites reinforced with ceramic fibers have attracted considerable interest because of their potential advantages over conventional materials in high-temperature load-bearing applications. Those advantages include enhanced strength and toughness and high resistance to corrosion and wear in severe service environment. Yet, their reliability under long-term sustained loading conditions remains to be quantitatively established so that lifetimes can be estimated to assure reliable service and prevent premature failure. Experimental observations on the microstructure of the creep-ruptured specimens indicate that the formation and propagation of microcracks along interfaces between fiber/matrix and matrix/matrix constitute major damage, and damage accumulation leading to final coalescence of these microcracks is responsible for creep fracture (Wiederhorn et al., 1988; Carroll et al., 1989; Hockey et al., 1991). This mode of rupture is in distinct contrast with the conventional short-term fracture mode wherein fiber bridging and fiber pull-out are the prevailing phenomena (Budiansky et al., 1986; Evans and McMeeking, 1986; Marshall and Cox, 1987; Nair et al., 1991). Thus, continuum damage mechanics approach rather than fracture mechanics is more appropriate in the description of the progressive damage state. Damage evolution law is then required for an accurate prediction of service life when it is controlled by slow crack growth. The present paper aims to provide a physics-based foundation for rational development of damage evolution laws.

Literature review indicate that because most of the microcracks are found at interfaces oriented in the direction normal to the principal tensile stress axis, the role that mass transport plays in crack growth and linkage must be an important aspect of the creep-rupture process. Diffusion-induced crack growth in single-phase isotropic media has been considered previously (Chuang, 1982). Strictly speaking, though, there exists no such a medium in reality even within monolithic systems, as the adjoining grains surrounding the crack are likely to possess anisotropic materials constants. For example, surface free energy and diffusivity are known to be dependent on the crystal orientation. The present work extends the treatment to a bimaterial system with a major goal of solving for the crack-tip velocity in the steady-state creep stage as a function of thermal-mechanical loading parameters.

At elevated temperatures in excess of about one third of the homologous temperature, mass transport becomes activated along high diffusivity paths such as interfaces and internal free surfaces. Because it takes less energy to form a void at an interface than in the bulk, creep cavities are predominantly observed at interfaces, rather than inside grains. Moreover, of all boundaries, cavitation seems to favor those normal to the principal tensile stress axis (cf. Fig. 1). These observations suggest that stress-driven diffusion around the cavity periphery plays an important role in the cavity growth process. One possible and convincing mechanism leading to cavity growth involves a coupled process of transportation of species along cavity surfaces towards the apex via surface diffusion and, from there, driving the atoms further away along the interface via grain-boundary diffusion. In this manner, the cavity tip is allowed to advance in a steady-state fashion. A direct proof showing this mechanism is indeed operative was provided by Varma and

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the International Mechanical Engineering Congress and Exposition, the Winter Annual Meeting of The ASME, Atlanta, GA, Nov. 17-22, 1996.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Apr. 21, 1995; final revision, Feb. 26, 1996. Associate Technical Editor: C. F. Shih.

Paper No. 96-WA/APM-7.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

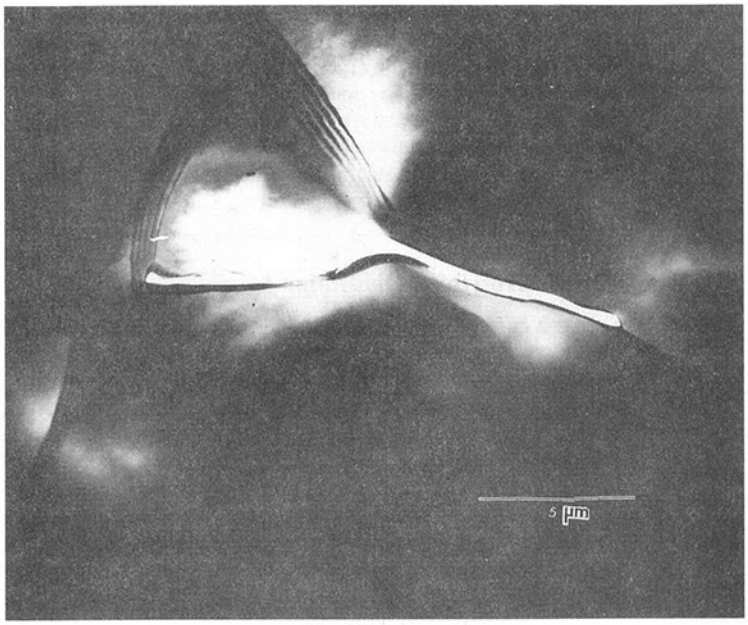


Fig. 1(a)

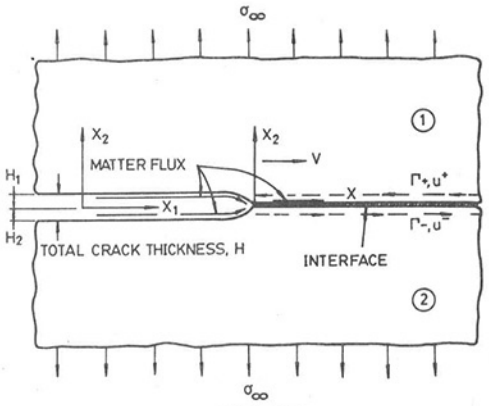


Fig. 1(b)

Fig. 1 Modeling diffusive crack growth: (*a*) TEM photo showing typical diffusional growth of a crack-like cavity. The microphoto was taken from the tensile side of a post-crept bend bar of a pure alumina tested at 1720°C, 20 MPa for 1800 s. (Chuang and Tighe, 1989). In the photo, the average grain size is 15 μm; total crack length about 17 μm normal to the tensile axis and crack thickness 125 nm. (*b*) Schematic sketch showing the proposed diffusive crack growth model. The crack is considered to travel from left to right at a constant speed V under a far-field tensile stress σ_∞ and an ambient temperature T. Phase 1 and 2 possess distinct properties. Single-phase system becomes the limiting case when the two phase properties are set to identical.

Dyson (1982) for a nickel-base alloy. Based on this specific growth kinetics, many diffusional *cavity* growth models have been proposed (Hull and Rimmer, 1959; Speight and Harris, 1967; Chuang and Rice, 1973; Raj and Ashby, 1975; Charles, 1976; Vitek, 1978; Chuang et al., 1979; Fuentes-Samaniego and Nix, 1981; Hirth, 1981; Svenson and Dunlop, 1981; Wilkinson and Vitek; 1982, Nix, 1983; Thouless and Liniger, 1995), of which the diffusional *crack* growth model is particularly relevant to the present case. Two separate systems will be discussed, namely, crack growth in a single-phase system and in a dissimilar or two-phase medium. The former aims at modeling matrix/matrix interfaces whereas the latter aims at fiber/matrix interfaces, taking the dissimilarity into account.

Crack Growth in Single-Phase Media

In order to acquire a final solution relating crack velocity to applied stress intensity, K, a precise analysis of the tensile stress distribution at the interface ahead of the moving crack-tip must be performed. In the absence of diffusion, the applied K induces the well-known elastic stress field with a characteristic $X^{-1/2}$ type singularity, where X is the distance away from the crack-tip. However, as matter diffuses from the crack surfaces and deposits along the interface, the so-called "wedging" effect is produced which alleviates the stress concentration at the crack-tip. Chuang (1982) has formulated this problem using the concept of infinitesimal edge dislocations to evaluate the residual stresses induced by mass transport along the interface. The result is an integral equation for the unknown stress distribution, $\sigma(x)$:

$$L^2 \cdot \sigma''(x) = \int_0^\infty \frac{\sigma(x_o)}{x - x_o}\, dx_o \qquad (1)$$

where the integration is performed in the sense of Cauchy principal value; the prime denotes d/dx and $x = X/L$ is a nondimensional boundary coordinate normalized by L:

$$L = \sqrt{\frac{\pi}{4} \cdot \frac{ED_b\delta_b\Omega}{(1 - \nu^2)VkT}} \qquad (2)$$

where E and ν are Young's modulus and Poisson's ratio, respectively; $D_b\delta_b$ is the grain-boundary diffusivity; Ω is the atomic volume, and V is the crack velocity. The magnitude of L is typically in the order of a micrometer (Chuang and Tighe, 1989). The initial conditions (in mathematical terms) for the tensile stress at the crack-tip are: $\sigma(0)/\sigma_o = 1$ and $\sigma'(0)L/\sigma_o = \alpha$, where α is a pure number depending on V (cf. Eq. (26) for a full expression).

A complete solution for the integral equation was obtained numerically (Chuang, 1982). It was found that the stress distribution is dependent on α which, in turn, is a function of V. Typical stress distributions for $\alpha = 0$, 0.25, 0.5, 0.75, 1.0, and 1.25, respectively, are plotted in Fig. 2. Here a hatch on the top

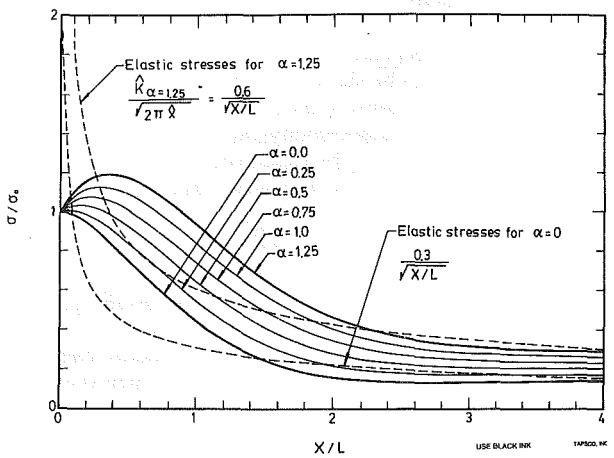


Fig. 2 Solutions of interfacial stress distribution ahead of the moving crack-tip. When L is defined properly, this plot is applicable to a general bimaterials system.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

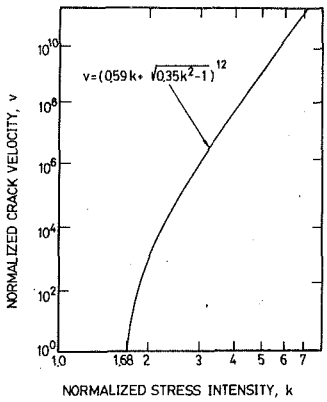


Fig. 3 Nondimensional *V-K* kinetic law predicted by the present diffusive crack growth theory at a bimaterial interface.

of a parameter denotes nondimensionalization. For the sake of comparison, the elastic stresses for the case of $\alpha = 1.25$ are given by a dashed line. It should be pointed out that $\alpha = 0.0$ and 1.25 represent two limiting cases, the former corresponding to $V =$ infinity whereas the latter corresponding to $V_{\min}$. Thus, any value of α falls outside the range becomes physically inadmissible. It is seen from this plot that the stress singularity at the crack-tip has been eliminated by diffusion and the peak stresses now occur at around $X \simeq 0.5L$ instead. The sizes of the diffusion zone are estimated to be approximately $4L$, beyond which the influence of diffusion becomes insignificant. The undisturbed far-field stresses are solely controlled by the applied K, which can then be related to $\sigma(0)$ and $\sigma'(0)$ at the crack-tip by the numerical solution. This relationship can be cast in the following form:

$$K = 0.75\sigma(0)L^{1/2} + 0.60\sigma'(0)L^{3/2}. \quad (3)$$

This is a significant result as this equation appears to have a widespread applicability, not only to the current model of diffusive crack growth in a single-phase elastic system, but also to a future model of bimaterial systems as will be discussed later. In addition, Cao et al. (1987) applied this equation to a problem of high-temperature stress-corrosion cracking of ceramics in which corrosive liquid product filled the crack-like cavity and served as a major species in the mass transport process.

The final goal of expressing K in terms of V, or vice versa, can be achieved by substituting the stress conditions at the crack-tip (namely, $\sigma(0) \propto V^{1/3}$; $\sigma'(0) \propto V^{2/3}$) in Eq. (3). The result is a V-K relationship in nondimensional form: $k = 0.845(v^{-1/12} + v^{1/12})$ where $k = K_I/K_G$ and $v = V/V_{\min}$. Here K_G[1] is the critical K-value based on the Griffith theory and $V_{\min}$ is a materials parameter depending on temperature, diffusivity, and elastic constants. It can be shown that if the physically inadmissible branch of the equation is discarded, a one-to-one relationship between the crack-tip velocity, V, and the applied stress intensity in mode I, K_I, can be established. Solving v in terms of k, one finally arrives at the following expression:

$$v = (0.59k + \sqrt{0.35k^2 - 1})^{12}. \quad (4)$$

It is seen that for a creep crack growing in a single-phase *elastic* material by diffusion, a unique V-K_I relationship emerges, irrespective of materials species or temperature. However, as indicated in Eq. (4), the V-K_I relationship, in general, cannot be cast in a simple form of a power-law equation which is conventionally adopted as an empirical expression. Figure 3 plots the equation in a double log space, and a few remarks are in order. First, there exists a threshold stress intensity, $K_{th} = 1.681K_G$, below which the applied stress is not sufficient to drive the crack and, as a result, the crack ceases to propagate and sintering may actually occur. Secondly, at higher K values, say $K > 4K_G$, a power-law equation, $V = (\text{constant}) \cdot K^{12}$ becomes an asymptote and is a fair representation of Eq. (4). Finally, for intermediate K-levels in the range $4K_G > K > 1.681K_G$, the theory predicts that the stress exponents vary from 12 to infinity.

[1] K_G is related to true surface energy and should not be confused with K_{IC} which is, in general, about one order of magnitude higher.

Crack Growth in Dissimilar Media

In the case of a crack growing along a fiber/matrix interface under creep conditions, the effects of dissimilarity, if any, on the crack growth behavior between the fiber and the matrix, in terms of differing physical and mechanical properties, becomes an interesting subject and deserves a thorough investigation. As the geometry dictates (see Fig. 1), the crack appears to be much thinner than the grain size or fiber diameter. Plane-strain conditions should prevail and this case can be modeled as a two-dimensional two-phase solid containing an interfacial crack, phase 1 representing the fiber, for example, and phase 2 the matrix. The problem now is to solve the crack growth rate for a given stress and temperature, and again this can be divided into two parts.

Crack-Tip Conditions. For the first part of the problem, Chuang et al. (1992) have recently investigated the crack-tip morphology that could be developed by surface diffusion-controlled crack growth along an interface between the two dissimilar phases. In contrast to the symmetric case developed in single-phase systems, four asymmetric cases are possible for two-phase media depending on the degree of dissimilarity in surface free energy and diffusivity of the two adjoining phases. Excluding the physically inadmissible case of the tip morphology, there are three possible cases of crack-tip morphology where the upper and lower cavity surfaces are no longer symmetrical with respect to the tip owing to differing properties of the dissimilar phases. A tip morphology map in the space of surface free energies, γ_1 versus γ_2, can be constructed to demonstrate prospective areas where each case applies. The near-tip shape again can be uniquely described by a logarithmic function as in the case of single-phase materials. However, the maximum half-thickness is now $2.0l$ instead of $1.41l$. Here l is a length parameter defined in Eq. (5). These results yield a more complex expression for the matter flux and the root radii (or the surface curvatures) at the crack-tip. These expressions involve physical properties of the two dissimilar phases. However, the velocity dependence remains the same, namely, the curvature at the tip, $\kappa_{\text{tip}} \sim V^{1/3}$, and the matter flux at the tip, $I_{\text{tip}} \sim V^{2/3}$. Accordingly, the boundary conditions at the moving crack-tip have the following relationships: $\sigma(0) \sim V^{1/3}$ and $\sigma'(0) \sim V^{2/3}$. Of course, the proportionality constants will have lengthy expressions in terms of materials constants and temperature. From Eq. (1), the stress solution $\sigma(x)$ or in tensorial term, $\sigma_{yy}(x)$ along the interface will relate the far-field stress to the normal stress at the moving crack-tip. In order to solve Eq. (1), initial conditions of stress at the crack tip have to be formulated. Expressions of the crack-tip stress and its first derivative in terms of the unknown, a priori, crack velocity, V, will allow the stress solution to yield the ultimate V-K_I relationship we desired.

The chemical potentials at the interface and at the crack surface can be expressed in the following forms, respectively: $\mu_b = -\sigma\Omega$ and $\mu_s = -\Omega\gamma\kappa$ where κ here is the surface curvature. The chemical potential at the crack-tip where the two crack surfaces joint the interface must have a unique value, otherwise there will be an unbounded flux there. This means that $\sigma_{\text{tip}} = -\gamma\kappa$. But $\kappa = -\sqrt{2(1-F)}/l$ where $F \equiv \sin\theta$ is the sine of the tip angle, and l is a length parameter defined by

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$l_i = \left(\frac{(D_s \delta_s \Omega)_i \gamma_i}{VkT} \right)^{1/3}. \quad (5)$$

Here D_s is surface diffusivity; δ_s, the thickness of the effective surface diffusion layer; the subscript i denotes phase i. Normally δ_s is in the order of $\Omega^{1/3}$. Let $\Delta_s = (D_{s2}\delta_2\Omega_2)/(D_{s1}\delta_1\Omega_1)$. It can then be shown (for derivation please see the Appendix and also Chuang et al., 1992) that the tip shape at the upper and lower surfaces has, respectively, the form

$$F_1 \equiv \sin(\theta_{\text{tip}})_1 = 1 - \frac{\hat{\gamma}_2^{1/3}(\hat{\gamma}_1 + \hat{\gamma}_2 - 1)}{\hat{\gamma}_1(\hat{\gamma}_2^{1/3} + \Delta_s^{2/3}\hat{\gamma}_1^{1/3})} \quad (6)$$

and

$$F_2 \equiv \sin(\theta_{\text{tip}})_2 = 1 - \frac{\hat{\gamma}_1^{1/3}\Delta_s^{2/3}(\hat{\gamma}_1 + \hat{\gamma}_2 - 1)}{\hat{\gamma}_2(\hat{\gamma}_2^{1/3} + \Delta_s^{2/3}\hat{\gamma}_1^{1/3})}, \quad (7)$$

where subscript 1 or 2 means phase 1 or 2, respectively. For nondimensional surface energy the normalizing parameter is γ_b. Since unique values for chemical potential must prevail at the crack-tip, $(\kappa_{\text{tip}})_i \gamma_i \Omega_i = \sigma_{\text{tip}}\Omega_i$ $(i = 1, 2)$. We have $\sigma_{\text{tip}} = -(\kappa_{\text{tip}})_i\gamma_i = \gamma_i\sqrt{2(1 - F_i)}/l_i$. Accordingly, from either Eq. (6) or (7), we finally arrive at the expression for the crack-tip stress, σ_{tip}, as a function of the crack velocity and other physical parameters of the two adjoining phases 1 and 2:

$$\sigma_{\text{tip}} = \sqrt{2}\sqrt{\frac{\gamma_1 + \gamma_2 - \gamma_b}{R_1 + R_2}}(VkT)^{1/3} \quad (8)$$

where R_1 and R_2 are defined as follows:

$$R_i \equiv \frac{(D_s\delta_s\Omega)_i^{2/3}}{\gamma_i^{1/3}}, \quad i = 1, 2. \quad (9)$$

Equation (8) shows that the stress at the crack tip is a function of crack-tip velocity V to the power of one third.

The first derivative of the crack-tip stress with respect to X, σ'_{tip} can be derived by a combination of grain-boundary diffusion equation, conservation of mass and steady-state conditions at $x = 0$. The result is

$$\sigma'_{\text{tip}} = \sqrt{2}\sqrt{(R_1 + R_2)(\gamma_1 + \gamma_2 - \gamma_b)}\frac{(VkT)^{2/3}}{D_b\delta_b\langle\Omega\rangle} \quad (10)$$

where $\langle\Omega\rangle$ is defined by the harmonic mean of Ω_1 and Ω_2 with a weighting factor R_1 and R_2, respectively:

$$\langle\Omega\rangle \equiv \frac{R_1 + R_2}{\frac{R_1}{\Omega_1} + \frac{R_2}{\Omega_2}} \quad (11)$$

and $D_b\delta_b\langle\Omega\rangle$ can be shown to be the sum of the grain-boundary diffusivities of the two respective phases, namely $D_b\delta_b\langle\Omega\rangle = (D_b\delta_b\Omega)_1 + (D_b\delta_b\Omega)_2$. In order to show this equality, the relationship of Eq. (A8) derived in the Appendix has been utilized. This is a very interesting result, because it shows that the effective atomic volume cannot be represented by a simple harmonic mean value; it must also involve parameters like surface diffusivity and surface free energy for each respective phase. However, in the event that the diffusing species are the same for both phases, $\Omega_1 = \Omega_2 = \Omega$, then according to Eq. (11), $\langle\Omega\rangle = \Omega$ even if $R_1 \neq R_2$. Equation (10) indicates that the first derivative of the crack-tip stress is in direct proportion to V to the power of two-thirds.

It is noteworthy that interchanging the subscript 1 and 2 in Eqs. (6) and (7) will alter the value of F's due to the asymmetric nature of tip morphology. However, values at the crack-tip should remain fixed when the top and bottom phases are interchanged. This can be verified by a mathematical operation of exchanging the subscript 1 and 2 in Eqs. (8) and (10). It can be seen that the results are invariant upon the exchanges.

Stress Analysis. To formulate the desired V-K_I relationship, a stress analysis at the interface ahead of the moving crack-tip must be carried out. The elastic stress field, in the absence of creep, in a bimaterial containing an interfacial crack has been investigated by many authors (Williams, 1959; England, 1965; Dunders, 1968; Rice, 1988; Suo, 1989). The results show many characteristics distinct from the single phase case: (1) Uniaxial stressing results in a mixed mode response. Thus, a far-field applied tensile stress will produce shear stresses at the interface. (2) The normal stresses at the interface ahead of the crack-tip induced by K_I show oscillatory behavior decaying with increasing distances from the crack-tip; and (3) The elastic crack-opening displacements show an unrealistic interpenetration behind the crack-tip, although the overlapping zone size is of atomic dimensions.

The above elastic behavior is indeed a fair representation of the short-term material responses. However, under long-term creep conditions where mass transport processes are active, it can be argued from physical principles that surface diffusion will prevent interpenetration and grain-boundary diffusion will relax the shear stresses produced at the interface resulting in grain-boundary sliding. This latter phenomenon is widely observed during high-temperature creep especially for ceramic systems where a liquid phase often exists between the grains. Based on the assumption that the interface ahead of the crack-tip cannot resist shear, it was found that mode I and II stress fields become decoupled. In addition, residual stresses induced by diffusion must be taken into account. To formulate the residual stress, consider an edge dislocation situated at an arbitrary location $x = x_o$ ahead of the semi-infinite crack-tip; then, the stress field generated by this dislocation has the form (England, 1965; Dunders, 1968; Rice, 1988; Suo, 1989):

$$\sigma_{yy} + i\sigma_{xy} = \sigma_\infty - \lambda(b_y + ib_x) \times \left[\frac{1}{x - x_o} + \pi i\beta\Delta(x - x_o) \right] \quad (12)$$

where $i = \sqrt{-1}$ is the unit of imaginary number, b_y and b_x are the Burgers vectors in the y and x directions, respectively, $\Delta(x)$ is the Dirac delta function, λ and β are elastic constants defined by

$$\lambda = \frac{\mu_1\mu_2[\mu_1(1 + \chi_2) + \mu_2(1 + \chi_1)]}{\pi(\mu_1 + \mu_2\chi_1)(\mu_2 + \mu_1\chi_2)}; \quad (13)$$

β is a Dundur's parameter (Dunders, 1968) defined by

$$\beta = \frac{\mu_1(\chi_2 - 1) - \mu_2(\chi_1 - 1)}{\mu_1(\chi_2 + 1) + \mu_2(\chi_1 + 1)} \quad (14)$$

where μ in these expressions is the shear modulus; and $\chi = 3 - 4\nu$ for plane-strain conditions, ν being Poisson's ratio. Equation (12) is a result with no crack-dislocation interactions under the remote tensile loading conditions. In addition to the normal stresses generated at the interface, shear stresses are produced there which would not be there under a symmetrical remote loading conditions *were* the system monolithic rather than dissimilar.

Let us consider an arbitrary distribution of dislocations along $y = 0$ under the action of a remote stress σ_∞ such that $db = \delta' dx$. The stresses at any point x can be integrated to yield the following forms:

$$\sigma_{yy}(x) = \sigma_\infty - \lambda\int_{-\infty}^{\infty} \frac{\delta'_y(x_o)}{x - x_o} dx_o + \lambda\pi\beta\delta'_x(x) \quad (15)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and

$$\sigma_{xy}(x) = -\lambda \int_{-\infty}^{\infty} \frac{\delta_x'(x_o)}{x - x_o}\, dx_o - \lambda\pi\beta\delta_y'(x). \tag{16}$$

Now, it may be argued that there exists a physical requirement that the interface cannot resist shear stresses (i.e., slipping interface) because of high-temperature creep. This is especially so in the case of advanced ceramics where flat boundaries are often coated with thin films of liquid phases that are introduced as sintering aids. This means that $\sigma_{xy} = 0$ everywhere along the interface ahead of the crack-tip. In addition, there is a relationship between δ_x and δ_y, namely,

$$\delta_y'(x) = \frac{-1}{\pi\beta} \int_{-\infty}^{\infty} \frac{\delta_x'(x_o)}{(x - x_o)}\, dx_o \tag{17}$$

and Hilbert transformation leads to

$$\delta_x'(x) = \frac{\beta}{\pi} \int_{-\infty}^{\infty} \frac{\delta_y'(x_o)}{(x - x_o)}\, dx_o. \tag{18}$$

Substituting Eqn. (18) into Eqn. (15), we have

$$\sigma_{yy}(x) = \sigma_\infty - \lambda(1 - \beta^2) \int_{-\infty}^{\infty} \frac{\delta_y'(x_o)}{(x - x_o)}\, dx_o \tag{19}$$

Again, by applying the Hilbert transformation, we finally arrive at

$$\pi^2\lambda(1 - \beta^2)\delta_y'(x) = \int_0^{\infty} \frac{\sigma_{yy}(x_o)}{x - x_o}\, dx_o \tag{20}$$

because in the traction-free crack plane ($y = 0, x < 0$) $\sigma_{yy} \equiv \sigma = 0$. Using grain-boundary diffusion equation, $I_b = D_b\delta_b/kT \cdot (d\sigma/dx)$, where $D_b\delta_b = (D_b\delta_b)_1 + (D_b\delta_b)_2$ and steady-state condition, $\delta_y(x) = [(I_b(x)\Omega)_1 + (I_b(x)\Omega)_2]/V = I_b(x)\langle\Omega\rangle/V$, we arrive at $\delta_y(x) = D_b\delta_b\langle\Omega\rangle/VkT \cdot (d\sigma/dx)$. Substituting in Eq. (20) to eliminate δ_y' yields the following integro-differential equation for the unknown $\sigma(x)$:

$$L^2 \cdot \sigma''(x) = \int_0^{\infty} \frac{\sigma(x_o)}{x - x_o}\, dx_o \tag{21}$$

where a new L is defined by

$$L = \sqrt{\frac{\pi}{4} \ll \frac{E}{1 - \nu^2} \gg \frac{D_b\delta_b\langle\Omega\rangle}{VkT}} \tag{22}$$

where $\langle\langle \cdot \rangle\rangle$ denotes simple harmonic mean of the two phases, and the equality of $\lambda(1 - \beta^2) = (2/\pi)\langle\langle \mu/(1 + \chi)\rangle\rangle = (1/4\pi)\langle\langle E/(1 - \nu^2)\rangle\rangle$ has been used. Note that this equality can be proved algebraically from the definition of λ and β given by Eqs. (13) and (14). It was found, after comparing Eq. (21) with (1) that the tensile creep stresses have the same solutions as indicated in Fig. 2, *if* a new length parameter, L, defined in Eq. (22) is used as a scaling parameter.

As a result, the V-K_I law for bimaterial systems has the same form as the single-phase systems depicted by Eq. (4), except that the normalizing parameters $V_{\min}$ and K_G will have different expressions involving the materials constants associated with the two-phase properties from Eqs. (8), (10), and (22). The results are

$$V_{\min} = 0.127\, \frac{(R_1 + R_2)^6}{kT} \left[\frac{\left\langle\left\langle \frac{E}{1 - \nu^2} \right\rangle\right\rangle}{D_b\delta_b\langle\Omega\rangle} \right]^3 \tag{23}$$

$$K_{th} = 1.681 K_G \tag{24}$$

where

$$K_G = \sqrt{(\gamma_1 + \gamma_2 - \gamma_b)\left\langle\left\langle \frac{E}{1 - \nu^2} \right\rangle\right\rangle} \tag{25}$$

is the stress intensity for the Griffith cracks in the interface between phase 1 and 2. Furthermore, a relationship between α, $V_{\min}$ and V is observed to exist, namely $V_{\min}/V = 0.262\alpha^6$, where α has the following expression in terms of V and other physical parameters:

$$\alpha = \sqrt{\frac{\pi}{4} \frac{\ll E/(1 - \nu^2) \gg}{D_b\delta_b\langle\Omega\rangle}} \frac{R_1 + R_2}{(kTV)^{1/6}}. \tag{26}$$

Since $V_{\min}$ by definition must be smaller than or equal to V, the range of α must lie between 0 and 1.25; i.e., $0 < \alpha \leq 1.25$. Figure 2 plots the normalized stress distributions for several values of α within this physically admissible range. In the case of α in excess of 1.25, higher crack velocities are predicted with decreasing applied stresses. This situation represents the physically unstable regime, and the corresponding branch of solution must be discarded. We also note from Fig. 2 that $\alpha = 0$ curve corresponds to $V = \infty$, and $\alpha = 1.25$ corresponds to $V = V_{\min}$. Thus, increasing α has an effect of reducing crack velocity. At a first glance, it looks as though higher stresses lead to lower crack velocities. This false impression is caused by plotting the stress distributions on the normalized physical plane. If one plots the stresses on the real physical plane, namely σ versus X, then the stresses for the $\alpha = 0$ curve will be dramatically enhanced, because σ_{tip} is much higher as it is proportional to $V^{1/3}$. As a result, the stresses at the far-field for the $\alpha = 0$ curve for the case of rapid crack propagation are actually higher than that for the $\alpha = 1.25$ curve where the crack velocity is minimum.

It should be noted that in the special case of a single-phase system where the properties of phase 1 are identical to that of phase 2, then $\beta = 0$ according to Eq. (14) and from Eq. (13) $\lambda = 2\mu/\pi(1 + \chi) = E/4\pi(1 - \nu^2)$; thus, L in Eq. (22) reduces to L of the single-phase case expressed in Eq. (2). In addition, all relevant parameters including $V_{\min}$, K_G, α, σ_{tip}, σ'_{tip} can be shown to be reduced to the corresponding expressions in the case of single-phase media. Because data on emerging two-phase materials are lacking, a direct comparison between the theory and experiment cannot be made at this time. However, preliminary data on the creep-rupture lifetime of silicon nitride reinforced with silicon carbide whiskers seemed to support the stress dependence on the crack-tip velocity (Chuang, 1992).

Energy Release Rate

It is well recognized in fracture mechanics that during crack growth some amount of energy is dissipated inside the body as a form of heat which is not recoverable. This released energy per unit increment area of cracked surface provides a driving "force" for crack extension. Here we attempt to investigate the sources of energy release rate accompanying diffusional crack growth along a bimaterials interface.

In the case of diffusional crack growth in single phase media, the energy release rate G during steady-state crack growth has been pursued by Chuang (1982, 1983) who found G consists of three terms, namely Griffith energy plus dissipated energy consumed in the mass transport process along the crack surfaces and the interface:

$$G \cdot V = \dot{F}_s + \dot{Q}_s + \dot{Q}_b \tag{27}$$

where F_s is the energy related to new surfaces created; Q_s and Q_b are the energy consumed by surface as well as grain-boundary diffusion; and a dot on the top denotes time rate. This equation assumes that the two phases behave elastically. Otherwise, any other extra active inelastic mechanisms such as dislocation

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

movement, which consume additional energy will have to be added to the right-hand side of Eq. (27) accordingly.

It is straightforward to derive F_s as $F_s = (\gamma_1 + \gamma_2 - \gamma_b)V$, since by growing a crack length da, an equal amount of da of new surface is created for phase 1 and 2, while losing an equal amount of interfacial area da. For the energy consumed by diffusion, it is in direct proportion to $\int \mathbf{I}\cdot(-\nabla\mu)ds$, since the inverse chemical potential gradient $(-\nabla\mu)$ represents the chemical force exerted on an atom, and $\mathbf{I}\cdot(-\nabla\mu)ds$ is therefore the total energy used locally to move atoms across ds per second. In terms of irreversal thermodynamics, the term represents local entropy production by matter flow at crack surfaces and interface. For surface diffusion, it can be shown that $Q_s = \int \mathbf{I}_1\cdot(-\nabla\mu_1)ds + \int \mathbf{I}_2\cdot(-\nabla\mu_2)ds = (\gamma_1 + \gamma_2 - \gamma_b)V$. Surprisingly, Q_s is identical to F_s and is independent of V. On the other hand, we found the heat consumed in interfacial diffusion is $Q_b = [J - 2(\gamma_1 + \gamma_2 - \gamma_b)]V$, where J is the well-known J-integral. Substituting these three items in Eq. (27), one obtains

$$G = J = -\int_0^\infty \sigma_{yy}\cdot(\partial\delta/\partial x)dx \tag{28}$$

where the J-integral is carried out along the paths $\Gamma-$ and $\Gamma+$ where $\delta = u^+ - u^-$ (see Fig. 1(b)), and $\delta = \delta_1 + \delta_2$ is the total thickening induced by species 1 and 2 at a location x along the interface. Note that although shear displacements are nonzero at the interface, the shear work, however, is zero due to vanishing shear stresses over there; thus contributes nothing to Eq. (28). The equation also neglects the strain energy contribution as the stress concentration at the crack tip has been alleviated by back stress produced by diffusion (see Fig. 2). The numerical solution given by Chuang (1982) indicated that this approximation produces less than five percent error. It should also be noted that a rigorous analysis on the energy release rate has been carried out which takes into account the strain energy contributions (Chuang, 1983). It was shown that G is exactly identical to J, if the definition of the J-integral in the context of *finite deformation* is adopted. Physically, this means that the energy release rate G stems from the mechanical work done by the interfacial stresses on opening up the interface (i.e., the so-called wedging effect), not from the loss of strain energy, as normally occurs in room-temperature elastic fracture.

Discussion

We have performed a rigorous analysis for a microcrack growing along an interface between two dissimilar media under creep conditions, and established the functional relationship between the crack velocity and the applied stress. We found the same relationship still applicable as in the case of crack growth in a single phase medium, *provided* a modified length scale be used wherein the harmonic means of elastic constants and atomic volumes of the two phases replace the single-valued ones.

We have also investigated the energy release rate associated with the diffusive crack growth along a bimaterial interface. We found that most of the energy release is in the form of heat loss during the mass transport process, and is produced by the mechanical work done by the interfacial stresses on opening of the interface for mass insertion. In sharp contrast with elastic fracture which is dominated by strain energy release, there is not much strain energy involved for high-temperature fracture as the creep stresses are much lower in comparison with elastic stresses where stress concentration at the crack-tip dominates the stress field.

During the course of the investigation, we also detected a constraint that must be imposed on the relationship between the diffusivities of the two phase species during the coupled process of mass transport. We found that the following relationship must exist between Δ_s and Δ_b in order to activate the coupled process:

$$\Delta_b = \left(\frac{\gamma_2}{\gamma_1}\right)^{-1/3}\Delta_s^{2/3} \tag{29}$$

where $\Delta_b \equiv (D_b\delta_b\Omega)_2/(D_b\delta_b\Omega)_1$ is the ratio of grain-boundary diffusivity of species 1 to that of species 2. A detailed derivation of Eq. (29) is given in the Appendix. For fixed values of surface free energies and surface diffusivities of the two species, the grain-boundary diffusivity of species 2 cannot have an arbitrary value, its value must be constrained and be related to the grain-boundary diffusivity of species 1 as dictated by Eq. (29): $(D_b\delta_b\Omega)_2 = (\gamma_2/\gamma_1)^{-1/3}\Delta_s^{2/3}(D_b\delta_b\Omega)_1$. The implication of this constraint on the microstructure that could be developed is very significant. Suppose that the value of grain-boundary diffusivity of species 2 deviated from the value given in Eq. (29). One possible scenario, as discussed in the Appendix, would be a development of degenerated tips for which $V_1 \neq V_2$, giving rise to a sliding-like cavity morphology as observed by Chen and Argon (1981) for type 304 stainless steel. If this is the case, the cavity-tip morphology may appear to be the consequence of grain-boundary sliding, instead, it may in fact be due to an incompatibility of gain boundary diffusivities of the two diffusing species: the "sliding" rate (i.e., V_2/V_1) is in direct proportion to this incompatibility to the power of three-halves, according to Eq. (A7) in the Appendix. Of course, once the degenerated tips are developed, a new analysis is required for the phase of slower moving tip because the interface ahead ceases to exist, and becomes a portion of the free surface. Nonetheless, the current analysis does suggest the propensity of this odd microstructural development. On the other hand, the observed sliding could be due to the dissimilar elastic properties of the adjoining phases, not from the applied shear stress-induced sliding.

It should also be noted that for perfect interfaces that can resist shear, a mix-mode stress field will be produced even for a pure mode I remote loading conditions. However, for slipping interfaces assumed in the present study, mode I and II are decoupled as the interfacial sliding will relieve the coupling effects. It is fair to say that in the case of perfect bonding at the interfaces, the solution will become more complicated, and we leave this interesting subject for future research.

The subject of high-temperature crack growth along a metal/ceramic interface under creep conditions has attracted much attentions in the area of surface modification engineering such as coatings and thin or thick film studies. Accordingly, it is interesting to compare the creep crack growth behaviors between the ceramic/metal systems and their metallic counterparts based on the present general theory. Let us designate phase 1 as metal and phase 2 as ceramic in a bimaterial system and subscript s as the same metal in a single-phase system. In general, at a fixed temperature and external loading conditions, the diffusivities of metals are much higher than that of the ceramics. Hence, we may assume that $(D_b\delta_b)_2 = (D_s\delta_s)_2 = 0$. This means that $R_2 = 0$ according to Eq. (9); $\langle\Omega\rangle = \Omega_1 = \Omega$ according to Eq. (11) and $D_b\delta_b\langle\Omega\rangle = D_b\delta_b\Omega$. Furthermore, since the Young's modulus of a ceramic is at least several times higher than that of a typical metal, we may assume E_2 approaches infinity. Thus, $\ll E/(1-\nu^2)\gg = 2E_1/(1-\nu^2)$ based on the definition of the harmonic means and the fact that the Poisson's ratios for both systems are approximately equal. Feeding this information to the formula and assuming $\gamma_1 = \gamma_2 = \gamma_s$, we find that the minimum crack velocity which is a material constant, $V_{\min}$ for the metal/ceramics system is predicted to be one-eighth of $V_{\min}$ for the single metallic system, namely $(V_{\min})_1 = \frac{1}{8}(V_{\min})_s$ according to Eq. (23). In addition, the Griffith K-levels have the following relationship: $(K_G)_1 = \sqrt{2}(K_G)_s$. Therefore, the theory predicts a higher threshold stress below which the crack will

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

not grow for the ceramic/metal bimaterial system than the single metallic system. Moreover, the minimum crack speed is predicted to be one order of magnitude lower for the ceramic/metal system. For the sake of illustration, suppose an identical external stress is applied which corresponds to $K_a = 1.681(K_G)_1 = 2.377(K_G)_s$ to both systems at the same temperature. Then $V_1 = \frac{1}{8}(V_{\min})_s$ and $V_s = 35{,}000(V_{\min})_s$ according to Eq. (4). We finally arrive at $V_s = 280{,}000V_1$ under the same applied stress and temperature. Hence, we can see a dramatic effect of the presence of a rigid phase on the crack growth behavior of a bimaterial system.

Conclusion

We have derived the V-K_I relationship for diffusion-induced microcrack growth in dissimilar media under sustained loads at elevated temperatures. By assuming the slipping interface that cannot resist shear, we found that the functional dependence of the crack-tip velocity on the applied K_I is similar to the case of single-phase media if a new length parameter is used. Furthermore, the two materials constants, $V_{\min}$ and K_G, must be modified by replacing the elastic constants with the harmonic means and atomic volume by the *effective* atomic volume for the two species involved in the mass transport process. At a fixed stress and temperature, it is predicted that the single metallic systems are more susceptible to crack growth than the counterpart of the ceramic/metal bimaterial systems. This crack-growth theory should be useful in aiding damage evolution law development for a reliable life prediction using a continuum damage mechanics approach.

As in the case of single-phase media, we found the energy release rate during steady-state diffusive crack growth derives mostly from the work done by interfacial stresses when the interface is thickened by the insertion of material which produces unrecoverable heat. In contrast with room temperature fracture, no strain energy release occurs for the case of long-term creep fracture.

During the course of the investigation, we also detected a new constraint that must be imposed on the grain-boundary diffusivities of the two species. Otherwise, peculiar phenomenon may result as a consequence, if the material fails to obey this condition.

Acknowledgments

It is a pleasure to thank NASA Lewis Research Center, High Temperature Materials Program, Structural Integrity Branch, Dr. John P. Gyekenyesi, for partial financial support under Interagency Order No. C-82000-R. We are also grateful to National Science Council of Taiwan for financial support.

References

Budiansky, B., Hutchinson, J. W., and Evans, A. G., 1986, "Matrix Fracture in Fiber Reinforced Ceramics," *Journal of the Mechanics and Physics of Solids,* Vol. 34, pp. 167–189.

Cao, H. C., Dalgleish, B. J., Hsueh, C.-H., and Evans, A. G., 1987, "High-Temperature Stress Corrosion Cracking in Ceramics," *Journal of the American Ceramic Society,* Vol. 70, pp. 257–264.

Carroll, D. F., Wiederhorn, S. M., and Roberts, D. E., 1989, "Technique for Tensile Creep Testing of Ceramics," *Journal of the American Ceramic Society,* Vol. 72, pp. 1610–1614.

Charles, R. J., 1976, "Diffusion-Controlled Stress Rupture of Polycrystalline Materials," *Metallurgical Transactions,* Vol. A7, pp. 1081–1089.

Chen, I. W., and Argon, A. S., 1981, "Creep Cavitation in 304 Stainless Steel," *Acta Metallurgica,* Vol. 29, pp. 1321–1333.

Chuang, T.-J., and Rice, J. R., 1973, "The Shape of Intergranular Creep Cracks Growing by Surface Diffusion," *Acta Metallurgica,* Vol. 21, pp. 1625–1628.

Chuang, T.-J., Kagawa, K. I., Rice, J. R., and Sills, B., 1979, "Overview No. 2: Non-equilibrium Models for Diffusive Cavitation of Grain Interfaces," ibid. Vol. 27, pp. 265–284.

Chuang, T.-J., 1982, "A Diffusive Crack Growth Model for Creep Fracture," *Journal of the American Ceramic Society,* Vol. 65, pp. 93–103.

Chuang, T.-J., 1983, "On the Energy Release Rate Associated with Diffusional Crack Growth," *International Journal of Fracture,* Vol. 23, pp. 229–242.

Chuang, T.-J., and Tighe, N. J., 1989, "Diffusional Crack Growth in Alumina," *Proceed. 3rd International Conference on Fundamental Fracture,* Irsee, Germany, pp. 129–132.

Chuang, T.-J., Chu, J. L., and Lee, S., 1992, "Asymmetric Tip Morphology of Creep Microcracks Growing along Bimaterials Interfaces," *Acta Metallurgica et Materialia,* Vol. 40, pp. 2683–2691.

Chuang, T.-J., 1992, "A Generic Model for Creep Rupture Lifetime Estimation on Fibrous Ceramic Composites," *Fracture Mechanics of Ceramics,* Vol. 10, R. C. Bradt, D. Hassleman, D. Munz, M. Sakai, and V. Shevchenko, eds. Plenum Press, New York, pp. 441–457.

Dundurs, J., 1968, "Elastic Interaction of Dislocations with Inhomogeneities," *Mathematical Theory of Dislocations,* ASME, New York, pp. 70–115.

England, A. H., 1965, "A Crack Between Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 400–402.

Evans, A. G., and McMeeking, R. M., 1986, "On the Toughening of Ceramics by Strong Reinforcements," *Acta Metallurgica,* Vol. 34, pp. 2435–2441.

Fuentes-Samaniego, R., and Nix, W. D., 1981, "Steady-state Diffusional Growth of an Axisymmetric Cavity on a Grain Boundary," *Philosophical Magazine,* Vol. 44, pp. 601–612.

Hirth, J. P., 1981, "Nucleation of Void Sheets in Creep and Hydrogen Attack," *Res. Mechanica Letters,* Vol. 1, pp. 3–5.

Hockey, B. J., Wiederhorn, S. M., Liu, W., Baldoni, J. G., and Buljan, S.-T., 1991, "Tensile Creep of Whisker-Reinforced Silicon Nitride," *Journal of Materials Science,* Vol. 26, pp. 3931–3939.

Hull, D., and Rimmer, D. E., 1959, "The Growth of Grain-Boundary Voids under Stress," *Philosophical Magazine,* Vol. 42, pp. 673–687.

Marshall, D. B., and Cox, B. N., 1987, "Tensile Fracture of Brittle Matrix Composites: Influence of Fiber Strength," *Acta Metallurgica,* Vol. 35, pp. 2607–2619.

Nair, S. V., Jakus, K., and Lardner, T. J., 1991, "The Mechanics of Matrix Cracking in Fiber Reinforced Ceramic Composites Containing a Viscous Interface," *Mechanics of Materials,* Vol. 12, pp. 229–244.

Nix, W. D., 1983, "Introduction to the Viewpoint Set on Creep Cavitation," *Scripta Metallurgica,* Vol. 17, pp. 1–43.

Raj, R., and Ashby, M. F., 1975, "Intergranular Fracture at Elevated Temperature," *Acta Metallurgica,* Vol. 23, pp. 653–666.

Rice, J. R., and Chuang, T.-J., 1981, "Energy Variations in Diffusive Cavity Growth," *Journal of the American Ceramic Society,* Vol. 64, pp. 46–53.

Rice, J. R., 1988, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 98–103.

Speight, M. V., and Harris, J. E., 1967, "Kinetics of Stress-Induced Growth of Grain-Boundary Voids," *Metal Science Journal,* Vol. 1, pp. 83–85.

Suo, Z., 1989, "Singularities Interacting with Interfaces and Cracks," *International Journal of Solids and Structures,* Vol. 25, pp. 1133–1142.

Svensson, L.-E., and Dunlop, G. L., 1981, "Growth of Intergranular Creep Cavities," *Int. Metals Reviews,* Vol. 2, pp. 109–131.

Thouless, M. D., and Liniger, W., 1995, "Effects of Surface and Boundary Diffusion on Void Growth," *Acta Metallurgica et Materialia,* Vol. 43, pp. 2494–2500.

Varma, R. K., and Dyson, B. F., 1982, "Metallographic Detection of Atom-plating due to Cavity Growth," *Scripta Metallurgica,* Vol. 16, pp. 1279–1284.

Vitek, V., 1978, "A Theory of Diffusion Controlled Intergranular Crack Growth," *Acta Metallurgica,* Vol. 26, pp. 1345–1356.

Wiederhorn, S. M., Roberts, D. E., Chuang, T.-J., and Chuck, L., 1988, "Damage-Enhanced Creep in a Siliconized Silicon Carbide: Phenomenology," *Journal of the American Ceramic Society,* Vol. 71, pp. 602–608.

Wilkinson, D. S., and Vitek, V., 1982, "The Propagation of Cracks by Cavitation: A General Theory," *Acta Metallurgica,* Vol. 30, pp. 1723–1733.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bull. Seismol. Soc. America,* Vol. 49, pp. 199–204.

APPENDIX

Derivation of the Constraint on Diffusivities, Eq. (29)

In the Discussion section, we have merely stated that there is a constraint imposed on the grain-boundary diffusivities of the two phases. Here, we wish to present a detailed derivation for this constraint. We begin with the fundamentals of the mass transport processes.

A1 Chemical Potentials. The formulation of boundary values of chemical potential at free surfaces and interfaces have been discussed in the literature (see, for example, Rice and Chuang, 1981). Basically they can be derived from the principles of thermodynamics, using Herring's procedure. The expression for chemical potential of an atom at a free surface has a form: $\mu_s = -\gamma_s \kappa \Omega$ when the flat surface at a free stress state is treated as a reference state and the strain energy contribution is neglected. However, after solving the near-tip cavity shape developed by surface diffusion (Chuang et al., 1992), the curvature of the cavity surface everywhere can be expressed as $\sqrt{2(1 - F_i)}/l_i$. Hence, for the bimaterial systems considered

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

here, we have $(\mu_s)_1 = -\gamma_1\kappa_1\Omega_1$ and $(\mu_s)_2 = -\gamma_2\kappa_2\Omega_2$. Now, the chemical potential per atom at an arbitrary location X along the interface is $(\mu_b)_1 = -\sigma\Omega_1$ for phase 1, if the stress-free state is regarded as a reference state. Similarly, $(\mu_b)_2 = -\sigma\Omega_2$ for phase 2. Here we assume the interface is a perfect matter sink (or vacancy source) so that vacancy concentration can always maintain its equilibrium level. Once again, the strain energy contributions are justifiably neglected.

A2 Matter Flux. According to the Fick's first law in the mass transport theory, gradient of chemical potentials leads to an inverse flux in vectorial sense, namely $I = \text{const.}\ (-\text{grad.}\ \mu)$. The version of Fick's law in the surface or grain-boundary diffusion can be cast in the form: $I = -D\delta/\Omega kT \cdot (d\mu/dx)$. After substituting the expressions for the chemical potentials, in the case of surface diffusion, the surface flux for phase i becomes

$$(I_s)_i = \frac{(D_s\delta_s\gamma)_i}{kT}\frac{d\kappa_i}{ds} \tag{A1}$$

and in the case of interfacial diffusion,

$$(I_b)_i = \frac{(D_b\delta_b)_i}{kT}\frac{d\sigma}{dx} \tag{A2}$$

Since there are two species flowing in the interface, the total flux is $I_b = (I_b)_1 + (I_b)_2$, and an effective grain-boundary diffusion equation can be defined as $I_b = (D_b\delta_b/kT)\cdot(d\sigma/dx)$ where $D_b\delta_b = (D_b\delta_b)_1 + (D_b\delta_b)_2$. Evaluating the surface fluxes and boundary fluxes for phase i at the crack tip, we have from Eq. (A1)

$$(\Omega \cdot I_s)_i = \sqrt{2(1 - F_i)}\, l_i V \tag{A3}$$

where l_i was given in Eq. (5) in the text. The boundary flux for phase i at the crack-tip has the form

$$(I_b)_i = \frac{(D_b\delta_b)_i}{kT}\sigma'_{\text{tip}}. \tag{A4}$$

A3 Continuity and Mass Conservation. So far we have derived the expressions for chemical potentials and matter fluxes for phase 1 and 2 at both the crack surfaces and at the interface, and in particular, at the crack-tip. To maintain a continuous flow and unique chemical potential at the crack-tip, and to satisfy the requirement that the mass conservation over there be fulfilled, we must have $(\mu_s)_i = (\mu_b)_i$ and $(I_s)_i = (I_b)_i$. Furthermore, the normal stress as well as its derivative with respect to x must have a unique value when evaluated from both phases. Accordingly, the requirement that the stress at the crack-tip must be unique leads to $(\mu_s/\Omega)_1 = (\mu_s/\Omega)_2$ or

$$\sqrt{\Delta_s}\gamma_1(1 - F_1)^{3/4} = \gamma_2(1 - F_2)^{3/4}. \tag{A5}$$

In addition, it has been shown that the equilibrium angles at the tip during cavity growth must satisfy the following equality (Rice and Chuang, 1981): $(\gamma_1 F_1 + \gamma_2 F_2 - \gamma_b)V = 0$ if the cavity apex does not constitute a "point source" of entropy production. Since $V \neq 0$ we have

$$\gamma_1 F_1 + \gamma_2 F_2 - \gamma_b = 0. \tag{A6}$$

Solving for the two unknowns F_1 and F_2 from Eqs. (A5) and (A6), the crack tip angels were obtained as given by Eqs. (6) and (7) in the main text.

Now, imposing the unique value of $(\sigma')_{\text{tip}}$ for both phases from Eq. (A4) and utilizing the expressions for F_i given in Eqs. (6)–(7), we find the ratio of the grain-boundary diffusivity of phase 1 to phase 2, $\Delta_b \equiv (D_b\delta_b\Omega)_2/(D_b\delta_b\Omega)_1$ must satisfy

$$\Delta_b = \left(\frac{R_2}{R_1}\right)\left(\frac{V_2}{V_1}\right)^{2/3} \tag{A7}$$

where V_i is the crack-tip velocity as measured from phase i and R_i was defined in Eq. (9) in the text. However, since there is only one crack-tip jointly shared by both phases, we must have $V_1 = V_2$, it follows that Eq. (A7) should reduce to $\Delta_b = R_2/R_1$ or

$$\Delta_b = \left(\frac{\gamma_2}{\gamma_1}\right)^{-1/3}\Delta_s^{2/3} \tag{A8}$$

as given in the Eq. (29) of the main text.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Sliding Velocity Flow of Rough Collisions in Multibody Systems

J. A. Batlle
Department of Mechanical Engineering,
Politechnical University of Catalunya,
Diagonal, 647;
08028 Barcelona, Spain

Single-point rough collisions in multibody systems with perfect constraints, under the assumptions of Coulomb's friction and infinite tangential stiffness at the collision point, require usually an integration over the normal impulse. The evolution of the sliding velocity, which is needed in the integration, is determined by an autonomous nonlinear flow. *The phase-space geometry of this flow depends upon five parameters associated with the system collision configuration and the friction coefficient μ, and gives a global picture of the system behavior in collisions with the configuration considered and arbitrary initial velocities. This geometry is studied using μ as a control parameter, and a set of threshold values of μ associated with changes in qualitative behavior are determined.*

1 Introduction

For smooth single-point collisions in multibody systems with perfect constraints, the use of Newton's or Poisson's rules leads to an "all linear" solution. However, this is not the case if friction is considered at the collision point, even if the assumption of infinite tangential stiffness at this point is made. There are two reasons for that: on one hand Newton's and Poisson's rules (with coefficient of restitution e, $0 \leq e \leq 1$, independent of the initial conditions) become energetically inconsistent and, on the other hand, the friction laws for forces cannot be extended to the friction impulses calculated over the collision interval if sliding stops or changes its direction. Consequently the equations of motion for rough collisions must be integrated over the collision interval. This integration is performed taking the normal impulse P_n as integrating variable instead of time.

Beghin (1951) was aware of this and addressed the problem of the three-dimensional collisions of two moving rough bodies by means of the Eulerian formulation, and used the work done by the normal force to determine the end of the collision. Energetical consistency was achieved by considering that a fraction of the compression work—from zero for perfectly plastic collisions, to one for perfectly ellastic ones—is restituted. This approach opened the door to the use of energetic coefficients of restitution.

More recently three-dimensional rough collisions have received new attention as a consequence of the mathematical modeling of multibody systems and the paradoxes derived from an erroneous application of Newton's and Poisson's rules in cases of energetical inconsistency. Batlle and Condomines (1991) extended Beghin's approach to multibody systems by means of the Lagrangian formulation, including the consideration of impulsive drivers—that guarantee the continuity of a set of generalized velocities during the collision—and an introductory study of the sliding velocity flow. More recently Batlle (1993, 1995) has developed a sufficiency condition for Newton's and Poisson's rules energetical consistency and has used the geometry of the sliding velocity flow to study the particular case of three-dimensional balanced collisions (which generalize the concept of collinear collisions). A thorough study of the sliding velocity flow is developed in this article. The Lagrange formulation was also used by Smith (1991) to obtain the equations of motion for colliding bodies with constraints. Stronge (1994) has addressed the case of two colliding bodies by using the Eulerian formulation and an energetic restitution coefficient e_w whose square defines the fraction of the compression work that is restituted. Bhatt and Koechling (1993) have made an introductory analysis of the sliding velocity flow for a body colliding with a rough fixed obstacle.

As it happens in system dynamics, a global understanding of the system behavior can be obtained from its phase-space geometry and this is usually more interesting than the calculation of particular solutions. For rough collisions in multibody systems, the geometry of the sliding velocity bidimensional phase-space allows not only a global understanding of the collision behavior but also the obtaining of threshold values of the friction coefficient μ. Both things are useful to design robust computer codes for percussive dynamics in multibody systems. The sliding velocity $\mathbf{v}_t$ at the collision point evolves according to an autonomous bidimensional first-order differential equation of motion or autonomous bidimensional flow. However, the end of the collision depends upon the evolution of the normal relative velocity v_n and is determined by the work done by the normal force. The evolution of the sliding direction $\mathbf{v}_t/|\mathbf{v}_t|$ is needed to integrate the normal relative velocity and the generalized velocities used to describe the degrees-of-freedom of the multibody system.

The flow of $\mathbf{v}_t$ depends upon the friction at the collision point. In this article it is shown that, under the assumption of Coulomb's law for the friction force and infinite tangential stiffness at the collision point, the $\mathbf{v}_t$ flow is completely defined by the friction coefficient μ and five parameters related to the system collision configuration. Several bifurcations—or qualitative changes in the dynamical behavior—occur as the friction coefficient μ increases from zero. An application example is presented. The results are easily extended to multibody systems with impulsive drivers (that is, drivers able to maintain the values of a set of generalized velocities during the collision interval).

2 The Lagrangian Formulation of Rough Collisions in Multibody Systems

The differential Lagrangian equations of motion for single-point rough collisions in multibody systems with perfect con-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Dec. 8, 1994; final revision, Oct. 3, 1995. Associate Technical Editor: E. J. Haug, Jr.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

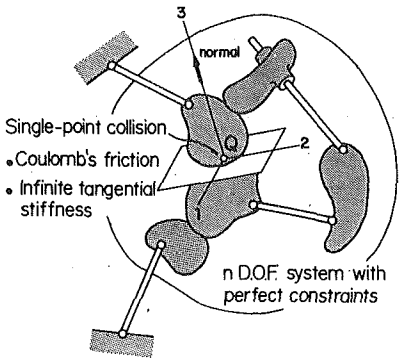


Fig. 1 Rough collision considered

straints and n degrees-of-freedom, described by means of the generalized velocities $\{u\}^T = \{u_1, u_2, \ldots, u_n\}$, (Fig. 1), can be written in a matrix form as

$$[M]\{du\} = \{d\Pi_n\} + \{d\Pi_t\} \equiv \{d\Pi\}, \tag{1}$$

where $[M]$ is the $n \times n$ inertia matrix of the system for the collision configuration and $\{d\Pi_n\}$ and $\{d\Pi_t\}$ are the generalized normal and tangential differential impulses, respectively. The n-vector $\{d\Pi\}$ is related to the differential normal and tangential impulses dP_n and $d\mathbf{P}_t$ by means of the equation

$$\{d\Pi\} = [k]\begin{Bmatrix} dP_n \\ d\mathbf{P}_t \end{Bmatrix} \equiv [k]\{dP\}, \tag{2}$$

where $[k]$ is a $n \times 3$ matrix of kinematical coefficients relating the normal velocity v_n and the tangential, or sliding, velocity $\mathbf{v}_t$ at the collision point to the generalized velocities:

$$\{v\} \equiv \begin{Bmatrix} v_n \\ \mathbf{v}_t \end{Bmatrix} = [\{\alpha\}[\beta]]^T\{u\} \equiv [k]^T\{u\}. \tag{3}$$

Equations (1), (2), and (3) lead to

$$\{du\} = [M]^{-1}[k]\{dP\}, \tag{4}$$

$$\{dv\} = [k]^T[M]^{-1}[k]\{dP\} \equiv [N]\{dP\}, \tag{5}$$

where $[N]$ is a 3×3 symmetrical positive definite matrix that is conveniently written in the form

$$[N] = \begin{bmatrix} \alpha^T M^{-1}\alpha & \alpha^T M^{-1}\beta \\ \beta^T M^{-1}\alpha & \beta^T M^{-1}\beta \end{bmatrix} \equiv \begin{bmatrix} a & \mathbf{h}^T \\ \mathbf{h} & [b] \end{bmatrix}. \tag{6}$$

The inertia operator $[N]$ depends only on the system configuration. It coincides with matrix $[M]$ used by Smith (1991).

In what follows Coulomb's friction and infinite tangential stiffness at the collision point are assumed.

When sliding occurs, $d\mathbf{P}_t$ can be expressed in terms of dP_n, the friction coefficient μ and the unit vector $\boldsymbol{\sigma} = \mathbf{v}_t/|\mathbf{v}_t|$ defining the sliding direction

$$d\mathbf{P}_t^s = -\mu\boldsymbol{\sigma}dP_n. \tag{7}$$

Substitution of (7) into (4) and (5) leads to

$$\{du^s\} = [M]^{-1}[k]\begin{Bmatrix} 1 \\ -\mu\boldsymbol{\sigma} \end{Bmatrix}dP_n \equiv \{C^s\}dP_n, \tag{8}$$

$$\begin{Bmatrix} dv_n^s \\ d\mathbf{v}_t \end{Bmatrix} = \begin{Bmatrix} a & -\mu\mathbf{h}^T\boldsymbol{\sigma} \\ \mathbf{h} & -\mu[b]\boldsymbol{\sigma} \end{Bmatrix}dP_n. \tag{9}$$

In the nonsliding case, the kinematical constraint $\mathbf{v}_t = 0$ leads, through (5), to

$$d\mathbf{P}_t^{ns} = -[b]^{-1}\mathbf{h}dP_n, \tag{10}$$

where, according to Coulomb's friction law, $|d\mathbf{P}_t^{ns}| \leq \mu dP_n$ and hence

$$|[b]^{-1}\mathbf{h}| \equiv \mu_c < \mu. \tag{11}$$

The value μ_c is the minimum friction coefficient that avoids the restart of sliding once it has stopped. It will be called *critical* friction coefficient.

The substitution of (10) into (4) and (5) leads to

$$\{du^{ns}\} = [M]^{-1}[k]\begin{Bmatrix} 1 \\ -[b]^{-1}\mathbf{h} \end{Bmatrix}dP_n \equiv \{C^{ns}\}dP_n, \tag{12}$$

$$\{dv_n^{ns}\} = (a - \mathbf{h}^T[b]^{-1}\mathbf{h})dP_n. \tag{13}$$

For systems with impulsive drivers, an analogous formulation is obtained, in which "a", $\mathbf{h}$ and $[b]$ have different expressions but keep the same properties (Batlle and Condomines, 1991; Batlle, 1993).

3 The $\mathbf{v}_t$ flow

Equation (9) shows that the evolution of $\mathbf{v}_t$ is defined by an autonomous flow described by a bidimensional nonlinear differential equation of first order

$$\frac{d\mathbf{v}_t}{dP_n} = \mathbf{h} - \mu[b]\frac{\mathbf{v}_t}{|\mathbf{v}_t|} \equiv \mathbf{h} - \mu[b]\boldsymbol{\sigma}, \tag{14}$$

where $\mathbf{h}$ and $[b]$ depend only on the system configuration. The vector $\mathbf{h}$ defines the rate of change of $\mathbf{v}_t$ with P_n in absence of friction, while $-\mu[b]\boldsymbol{\sigma}$ defines the rate associated with friction, which is strictly opposed to $\boldsymbol{\sigma}$ for the eigendirections of $[b]$. The solutions of (14), called hodographs of $\mathbf{v}_t$, cannot intersect each other. The set of hodographs generated by *all* the initial conditions (which is roughly analogous to that of the paths followed by the particles of a flowing fluid, and so justifies the word *flow* for Eq. (14)) gives a clear image of the system behavior for collisions with the same configuration but different initial velocities.

From the initial sliding velocity, the span along the flow line depends on the end of the collision, which is defined by the work done by the normal force

$$W_n = \int_0^{P_n} v_n dP_n = (1 - e_w^2)W_{\text{comp}}, \tag{15}$$

where W_{comp} is the compression work, or maximum negative value of W_n, and e_w, $0 \leq e_w \leq 1$, is a restitution factor associated with W_n (Batlle, 1993; Stronge, 1994).

As shown by (8) and (9), the evolution of v_n and $\{u\}$ depends upon the sliding direction $\boldsymbol{\sigma} = \mathbf{v}_t/|\mathbf{v}_t|$.

Equation (14) shows that the isoclines of $\mathbf{v}_t$ flow are radial lines from the origin and consequently *the flow is self-similar for changes on v_{t1}, and v_{t2} scales.* Among the isoclines, the asymptotes or invariant directions (those whose direction coincides with that of $d\mathbf{v}_t$ and, consequently, with that of $\mathbf{v}_t$ as well) and those orthogonal to the hodographs (the *flow change directions* in the Bhatt and Koechling work (1993)) are particularly important. Figure 2 defines the axes and notation used to describe the flow phase space geometry.

The orthogonal isoclines are defined by the unit vectors $\boldsymbol{\sigma}_\perp$ verifying $\boldsymbol{\sigma}_\perp^T d\mathbf{v}_t = 0$. According to (9)

$$\boldsymbol{\sigma}_\perp^T\mathbf{h} - \mu\boldsymbol{\sigma}_\perp^T[b]\boldsymbol{\sigma}_\perp = 0. \tag{16}$$

To determine the asymptotes, it is useful to introduce the transversal unit vector $\boldsymbol{\sigma}_T$, which is orthogonal to the isocline, defined by

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

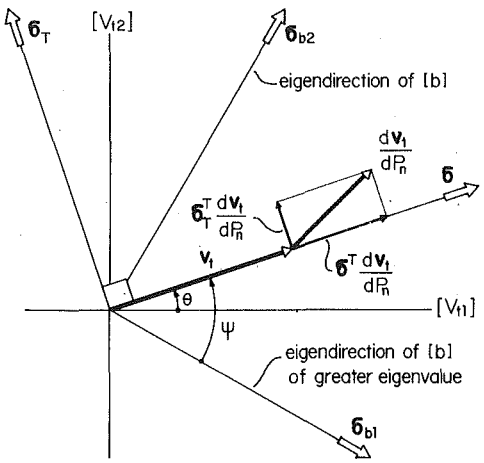


Fig. 2 Axes and notation used to describe the phase-space geometry of the v_t flow

$$\boldsymbol{\sigma}_T = [T]\boldsymbol{\sigma}; \quad [T] \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{17}$$

The asymptotes are defined by the unit vectors $\boldsymbol{\sigma}_A$ verifying

$$\boldsymbol{\sigma}_{AT}^T \mathbf{h} - \mu \boldsymbol{\sigma}_{AT}^T [b] \boldsymbol{\sigma}_A = 0. \tag{18}$$

4 The Critical Isocline

Both (16) and (18) are verified by the *critical isocline* $\boldsymbol{\sigma}_c$, for which $d\mathbf{v}_t = \mathbf{0}$. Equations (9) and (11) lead to

$$\boldsymbol{\sigma}_c = [b]^{-1}\mathbf{h}/\mu_c. \tag{19}$$

The required friction coefficient is equal to the *critical coefficient* μ_c defined by (11), above which sliding does not restart once it has stopped.

If b_1, b_2 (with $b_1 > b_2$) are the eigenvalues of $[b]$ and h_1, h_2 are the projections of $\mathbf{h}$ on the eigendirections $\boldsymbol{\sigma}_{b1}$, $\boldsymbol{\sigma}_{b2}$ (where $\boldsymbol{\sigma}_{b1}$ is taken to have positive projection of $\mathbf{h}$) then the angle ψ_c from the eigendirection $\boldsymbol{\sigma}_{b1}$ to the critical isocline is given by

$$\tan \psi_c = (b_1 \tilde{h}_2)/(b_2 \tilde{h}_1). \tag{20}$$

The hodographs starting at the critical isocline reduce to a single point. The other hodographs approach asymptotically the critical isocline with an angle γ determined by

$$\tan \gamma = \lim_{\boldsymbol{\sigma} \to \boldsymbol{\sigma}_c} \frac{\boldsymbol{\sigma}_T^T d\mathbf{v}_t}{\boldsymbol{\sigma}^T d\mathbf{v}_t} = \frac{\mathbf{h}^T [b^I] \mathbf{h}}{\mathbf{h}^T [b]^{-1} [T] \mathbf{h}}, \tag{21}$$

where

$$[b^I] \equiv [b]^{-1}[T]^T[b][T][b]^{-1}. \tag{22}$$

From (20) and (21)

$$\tan \gamma = -[b_1 \tan \psi_c + b_2 \cot \psi_c]/(b_1 - b_2). \tag{23}$$

5 Polar Plots for Friction Coefficient μ

In order to study the dependence of the orthogonal isoclines and the asymptotes upon the friction coefficient μ, it is interesting to introduce the plane of vectors $\boldsymbol{\mu} \equiv \mu\boldsymbol{\sigma}$ and plot the curves $\boldsymbol{\mu}_\perp(\boldsymbol{\sigma}) \equiv \mu_\perp \boldsymbol{\sigma}$ and $\boldsymbol{\mu}_A(\boldsymbol{\sigma}) \equiv \mu_A \boldsymbol{\sigma}$ corresponding to the values $\mu_\perp$ and μ_A for which a direction $\boldsymbol{\sigma}$ is an orthogonal isocline or an asymptote, respectively.

For the orthogonal isoclines, (16) leads to

$$\boldsymbol{\mu}_\perp(\boldsymbol{\sigma}) \equiv \mu_\perp \boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\sigma}^T \mathbf{h})/(\boldsymbol{\sigma}^T [b] \boldsymbol{\sigma}). \tag{24}$$

Equation (24) defines an ellipse of center at $\boldsymbol{\mu}_c/2$ and matrix $[b]$. Its axes are the eigendirections $\boldsymbol{\sigma}_{b1}$ and $\boldsymbol{\sigma}_{b2}$ of $[b]$ and it contains the origin and the point $\boldsymbol{\mu}_c \equiv \mu_c \boldsymbol{\sigma}_c$. In polar coordinates with polar axis $\boldsymbol{\sigma}_{b1}$,

$$\mu_\perp = (\tilde{h}_1 \cos \psi + \tilde{h}_2 \sin \psi)/(b_1 \cos^2 \psi + b_2 \sin^2 \psi). \tag{25}$$

For the asymptotes, (18) leads to

$$\boldsymbol{\mu}_A(\boldsymbol{\sigma}) \equiv \mu_A \boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\sigma}_T^T \mathbf{h})/(\boldsymbol{\sigma}_T^T [b] \boldsymbol{\sigma}). \tag{26}$$

Equation (26) defines an hyperbola of center $\boldsymbol{\mu}_{hs}/2$, being $\boldsymbol{\mu}_{hs}$ the vector $\boldsymbol{\mu}_A$ for a direction $\boldsymbol{\sigma}_{hs}$ symmetrical to $\boldsymbol{\sigma}_h \equiv \mathbf{h}/|\mathbf{h}|$ with respect to eigendirection $\boldsymbol{\sigma}_{b1}$:

$$\boldsymbol{\sigma}_{hs} = (\boldsymbol{\sigma}_{b1}\boldsymbol{\sigma}_{b1}^T - \boldsymbol{\sigma}_{b2}\boldsymbol{\sigma}_{b2}^T)\boldsymbol{\sigma}_h. \tag{27}$$

The hyperbola asymptotes are parallel to the eigendirections of $[b]$. In polar coordinates with polar axis $\boldsymbol{\sigma}_{b1}$,

$$\mu_A = (\tilde{h}_1 \sin \psi - \tilde{h}_2 \cos \psi)/(b_1 - b_2) \sin \psi \cos \psi). \tag{28}$$

The ellipse $\boldsymbol{\mu}_\perp(\boldsymbol{\sigma})$ and the hyperbola $\boldsymbol{\mu}_A(\boldsymbol{\sigma})$ for the application example presented in Section 9 are shown in Fig. 3.

The ellipse $\boldsymbol{\mu}_\perp(\boldsymbol{\sigma})$ and the hyperbola $\boldsymbol{\mu}_A(\boldsymbol{\sigma})$ intersect each other at the origin, orthogonally, and at $\boldsymbol{\mu}_c$. As $\mathbf{h} = \boldsymbol{\mu}_c[b]\boldsymbol{\sigma}_c$, the direction $\boldsymbol{\sigma}_h$ is the ellipse polar direction of $\boldsymbol{\sigma}_c$, that is, $\boldsymbol{\sigma}_h$ and the tangent to the ellipse at $\boldsymbol{\mu}_c$ are orthogonal. For a given value of μ, the intersections of a centered circumference of radius μ with the ellipse define the orthogonal isoclines, while the intersections with the hyperbola define the asymptotes.

6 Bifurcations Associated With Orthogonal Isoclines

For $\mu = 0$ there are two orthogonal isoclines which are tangent to the ellipse at the origin, and consequently are orthogonal to $\boldsymbol{\sigma}_h$ (the hodographs reduce in this case to straight lines parallel to $\mathbf{h}$). As μ increases, the orthogonal isoclines tilt towards the $\mathbf{h}$ side. According to the particular elements of $[b]$ and $\mathbf{h}$ the maximum number of intersections may be two or four, pairs of which may coincide in tangential intersections. This is illustrated in Fig. 4 for the application example of Section 9. In this case there are three bifurcation values of μ concerning the orthogonal isoclines. As μ increases from zero, there is a first tangential intersection for $\mu_{\perp 1}^*$ that originates two new orthogonal isoclines. These isoclines separate from each other as μ increases. For $\mu_{\perp 2}^*$, two orthogonal isoclines merge and disap-

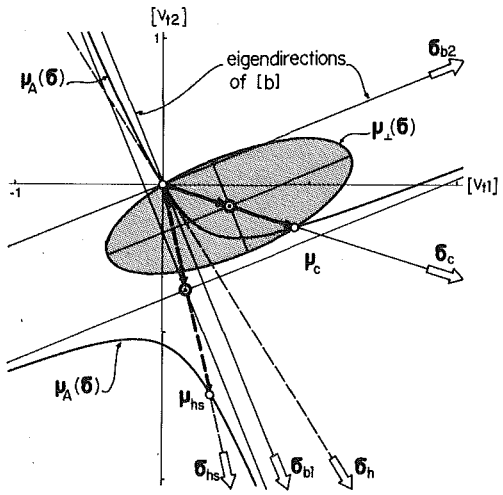


Fig. 3 Geometry of polar plots for $\mu_\perp(\sigma)$ and $\mu_A(\sigma)$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

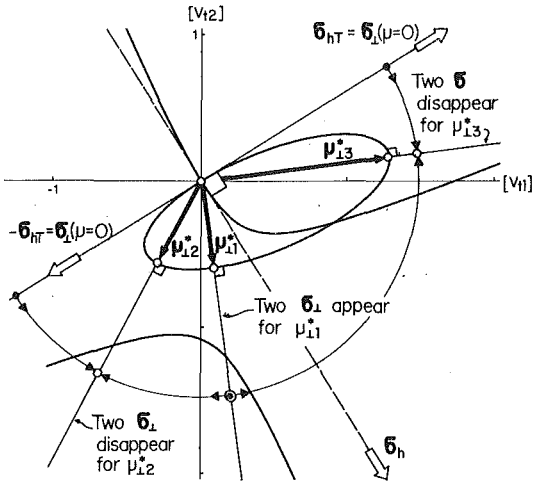


Fig. 4 Bifurcation vectors $\mu^*_{\perp 1}$, $\mu^*_{\perp 2}$, $\mu^*_{\perp 3}$ for the orthogonal isoclines. Evolution of the orthogonal isoclines for increasing μ.

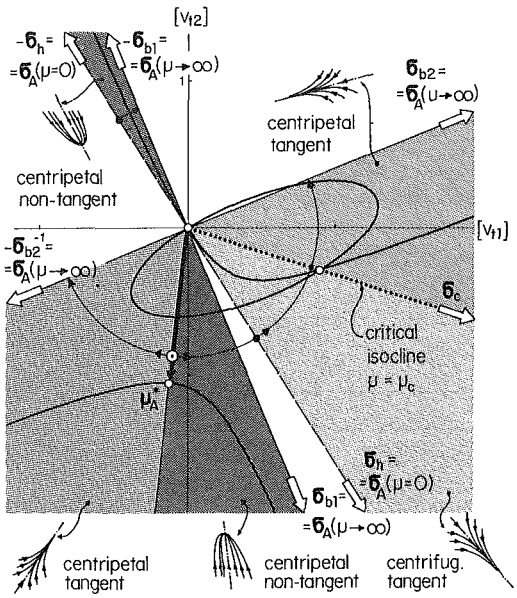


Fig. 5 Bifurcation vector μ^*_A for the asymptotes. Evolution of the asymptotes for increasing μ.

pear at a second tangential intersection. For $\mu^*_{\perp 3}$, the remaining two orthogonal isoclines merge and disappear. For higher values of μ there are no orthogonal isoclines. In cases where the maximum number of intersections between the ellipse $\boldsymbol{\mu}_\perp(\boldsymbol{\sigma})$ and the centered circumferences of radius μ is two, there is only the bifurcation associated with $\mu^*_{\perp 3}$. For it the two initial orthogonal isoclines merge and disappear.

For any value of $\mu < \mu^*_{\perp 3}$, the orthogonal isoclines separate sectors of different sign for radial projection $\boldsymbol{\sigma}^T d\mathbf{v}_t$. In sectors for which $\boldsymbol{\sigma}^T d\mathbf{v}_t > 0$, the hodographs have a centrifugal evolution while for $\boldsymbol{\sigma}^T d\mathbf{v}_t < 0$ their evolution is centripetal. For $\mu > \mu^*_{\perp 3}$, the hodographs have a centripetal evolution in all their points.

The directions $\boldsymbol{\sigma}^*_{\perp 1}$, $\boldsymbol{\sigma}^*_{\perp 2}$, and $\boldsymbol{\sigma}^*_{\perp 3}$ of the tangential bifurcation for orthogonal isoclines can be calculated from $\partial \mu_\perp / \partial \psi = 0$. This leads to equation

$$\tan^3 \psi^*_\perp - \frac{\tilde{h}_2}{\tilde{h}_1}\left(\frac{2b_1}{b_2} - 1\right)\tan^2 \psi^*_\perp + \left(2 - \frac{b_1}{b_2}\right)\tan \psi^*_\perp - \frac{\tilde{h}_2}{\tilde{h}_1}\frac{b_1}{b_2} = 0. \quad (29)$$

There is at least one solution of (29), the one corresponding to $\mu^*_{\perp 3}$, and there can be two other solutions: those corresponding to $\mu^*_{\perp 1}$ and $\mu^*_{\perp 2}$.

7 Bifurcation Associated With the Asymptotes

For $\mu = 0$ there are two asymptotes, a centrifugal one in the direction of $\mathbf{h}$ and one opposite to it that is centripetal. As μ increases, the centripetal asymptote tends towards $-\boldsymbol{\sigma}_{b1}$ while the centrifugal one tilts towards the critical isocline, and coincides with it for $\mu = \mu_c$. For $\mu > \mu_c$ this asymptote becomes centripetal and tends to the $\boldsymbol{\sigma}_{b2}$ direction as shown in Fig. 5 for the application example of Section 9. There is a single bifurcation value of μ concerning the asymptotes. There is always a certain value μ^*_A for which there is a tangent intersection with the other branch of the hyperbola that generates two new asymptotes which are centripetal. For $\mu > \mu^*_A$ these asymptotes will be determined by points of this branch of the hyperbola and tend, for $\mu \to \infty$, to the eigendirections of $[b]$. The direction $\boldsymbol{\sigma}^*_A$ of the tangential bifurcation for the asymptotes and μ^*_A can be found analytically from $\partial \mu_A / \partial \psi = 0$, which leads to

$$\tan \psi^*_A = -(\tilde{h}_2/\tilde{h}_1)^{1/3}, \quad (30)$$

$$\mu^*_A = [1 + (\tilde{h}_2/\tilde{h}_1)^{2/3}]^{3/2}\tilde{h}_1/(b_1 - b_2). \quad (31)$$

The asymptotes are alternatively of the *tangent* type and of the *nontangent* type (Fig. 5). The asymptotes tending to $\pm\boldsymbol{\sigma}_{b1}$ are of nontangent type and the asymptotes tending to $\pm\boldsymbol{\sigma}_{b2}$ are of tangent type.

For $\mathbf{h} = \mathbf{0}$, the asymptotes are the eigendirections of $[b]$ and consequently they are two eigendirections of the inertia operator $[N]$. This is the case of the balanced collisions (Batlle, 1996).

8 The Newtonian Asymptote

As it has been mentioned in the Introduction, Newton's rule (with coefficient of restitution e, $0 \le e \le 1$, independent from the initial conditions) is in general energetically inconsistent. However, there is a wide sufficiency condition that guarantees its energetical consistency (Batlle, 1993) which calls for a constant rate dv_n/dP_n throughout the collision. Whenever sliding is present this condition requires a constant sliding direction, so the hodograph must be an asymptote and it must not reach the origin, or just reach it at the end of the collision. If sliding stops and does not restart (so that there is a fraction of the collision interval without sliding), the sufficiency condition is

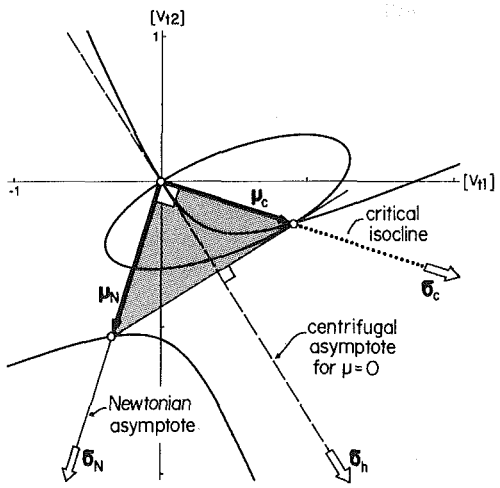


Fig. 6 Geometry of the Newtonian asymptote

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

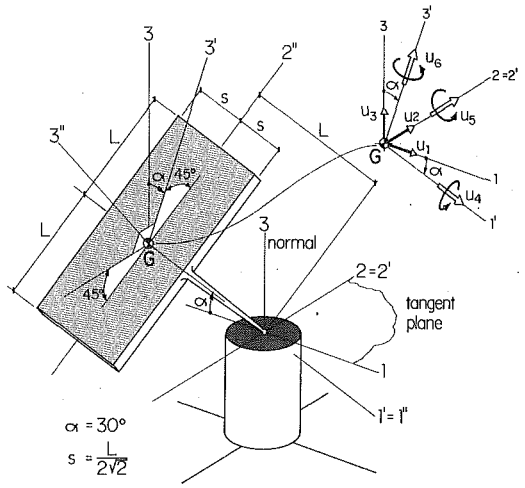


Fig. 7 Application example parameters and generalized velocities

not fulfilled in principle because dv_n/dP_n has a different value under sliding and nonsliding conditions. However, (9) and (10) show that dv_n/dP_n may have the same value, regardless of the sliding or nonsliding conditions, for a particular sliding direction $\boldsymbol{\sigma}_N$ and a particular value of $\mu = \mu_N$ verifying

$$\mu_N \mathbf{h}^T \boldsymbol{\sigma}_N = \mathbf{h}^T [b]^{-1} \mathbf{h}. \tag{32}$$

But $[b]^{-1}\mathbf{h} = \boldsymbol{\mu}_c$, and consequently

$$\mathbf{h}^T \boldsymbol{\mu}_N = \mathbf{h}^T \boldsymbol{\mu}_c, \tag{33}$$

according to (33), the vectors $\boldsymbol{\mu}_N$ and $\boldsymbol{\mu}_c$ have the same projection on $\mathbf{h}$. Furthermore, it is easy to check that $\boldsymbol{\mu}_N$ and $\boldsymbol{\mu}_c$ are orthogonal as shown in Fig. 6 for the application example of Section 9. (It is worth noting that this orthogonality is a general property of the hyperbola: any straight line perpendicular to the straight line tangent to the hyperbola at an arbitrary point P, intersects the hyperbola at points Q and Q' such that $\overline{\mathbf{PQ}}$ and $\overline{\mathbf{PQ}}'$ are orthogonal). The asymptote $\boldsymbol{\sigma}_N$, for which Newton's rule is energetically consistent even if the sliding stops will be called *Newtonian asymptote.*

9 Application Example

The single point collision of a rigid body with a rough fixed surface is considered. The mechanical and geometrical parameters at the collision configuration and the generalized velocities used are shown in Fig. 7. The mass of the rod from P to the rectangular plate is negligible. The analytical expressions for "a", $\mathbf{h}$ and $[b]$ will be found for an arbitrary angle α and for a rigid body with central inertia eigendirections 1", 2", 3" with principal inertia moments $m\rho^2$, $2m\rho^2\lambda_2$ and $2m\rho^2\lambda_3$, respectively. The inertia and kinematical matrices are

$$[M] = m\rho^2 \begin{bmatrix} \rho^{-2} & & & & & \\ & \rho^{-2} & & & & \\ & & \rho^{-2} & & & \\ & & & 1 & 0 & 0 \\ & & & 0 & \lambda_2+\lambda_3 & \lambda_2-\lambda_3 \\ & & & 0 & \lambda_2-\lambda_3 & \lambda_2+\lambda_3 \end{bmatrix},$$

$$[K]^T = \begin{bmatrix} 0 & 0 & 1 & 0 & -L\cos\alpha & 0 \\ 1 & 0 & 0 & 0 & -L\sin\alpha & 0 \\ 0 & 1 & 0 & 0 & 0 & L \end{bmatrix},$$

from which, using axes 1, 2 for $\mathbf{h}$ and $[b]$

$$a = \frac{1}{m}\left[1 + \left(\frac{L}{\rho}\right)^2 \frac{\lambda_2+\lambda_3}{4\lambda_2\lambda_3}\cos^2\alpha\right],$$

$$\mathbf{h} = \frac{1}{m}\begin{Bmatrix} \left(\frac{L}{\rho}\right)^2 \frac{\lambda_2+\lambda_3}{4\lambda_2\lambda_3}\sin\alpha\cos\alpha \\ -\left(\frac{L}{\rho}\right)^2 \frac{\lambda_3-\lambda_2}{4\lambda_2\lambda_3}\cos\alpha \end{Bmatrix},$$

$$[b] = \frac{1}{m}\begin{bmatrix} 1+\left(\frac{L}{\rho}\right)^2 \frac{\lambda_3-\lambda_2}{4\lambda_2\lambda_3}\sin^2\alpha & -\left(\frac{L}{\rho}\right)^2 \frac{\lambda_3-\lambda_2}{4\lambda_2\lambda_3}\sin\alpha \\ -\left(\frac{L}{\rho}\right)^2 \frac{\lambda_3-\lambda_2}{4\lambda_2\lambda_3}\sin\alpha & 1+\left(\frac{L}{\rho}\right)^2 \frac{\lambda_2+\lambda_3}{4\lambda_2\lambda_3} \end{bmatrix}.$$

For $\alpha = 30$ deg and the rectangular plate shown in Fig. 8, for which $\rho = L/\sqrt{3}$, $\lambda_2 = 1/8$ and $\lambda_3 = 9/8$,

$$a = m^{-1}6,$$

$$\mathbf{h} = (m^{-1}/\sqrt{3})\begin{Bmatrix} 5 \\ -8 \end{Bmatrix}; \quad \theta_h = -57.995°, \quad |\mathbf{h}| = 5.447,$$

$$[b] = (m^{-1}/3)\begin{bmatrix} 7 & -8 \\ -8 & 23 \end{bmatrix}.$$

The eigenvalues and eigendirections of $[b]$ are

$$b_1 = 8.771; \quad \boldsymbol{\sigma}_{b1}^T = \{0.383, -0.924\}; \quad \theta_{b1} = -67.500 \text{ deg},$$

$$b_2 = 1.229; \quad \boldsymbol{\sigma}_{b2}^T = \{0.924, 0.383\}; \quad \theta_{b1} = 22.500 \text{ deg}.$$

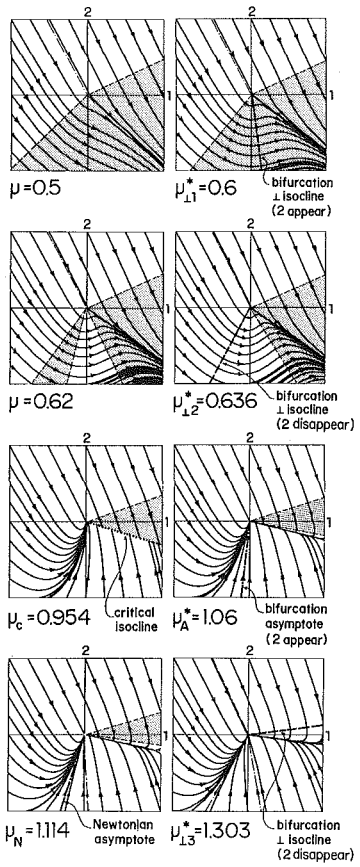


Fig. 8 Hodographs in the phase-plane of v_t for several values of μ. (hatched area: centrifugal evolution; nonhatched area: centripetal evolution)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The projections of $\mathbf{h}$ on the eigendirections $\boldsymbol{\sigma}_{b1}$ and $\boldsymbol{\sigma}_{b2}$ are

$$\tilde{h}_1 = m^{-1}5.372; \quad \tilde{h}_2 = m^{-1}0.899.$$

The critical $\boldsymbol{\mu}_c$ for the critical isocline, in polar notation is

$$\psi_c = 50.082 \text{ deg}; \quad \mu_c = 0.954,$$

and the angle γ of incidence of the hodographs on the critical isocline is $\gamma = -56.765$ deg.

The bifurcation vectors $\boldsymbol{\mu}_\perp^*$ for orthogonal isoclines are

$$\psi_{\perp 1}^* = -14.961 \text{ deg}; \quad \mu_{\perp 1}^* = 0.600,$$

$$\psi_{\perp 2}^* = -50.407 \text{ deg}; \quad \mu_{\perp 2}^* = 0.6367,$$

$$\psi_{\perp 3}^* = 74.873 \text{ deg}; \quad \mu_{\perp 3}^* = 1.303.$$

The bifurcation vector $\boldsymbol{\mu}_A^*$ for asymptotes is

$$\psi_A^* = -28.862 \text{ deg}; \quad \mu_A^* = 1.060.$$

The Newtonian asymptote corresponds to

$$\psi_N = -39.918 \text{ deg}; \quad \mu_N = 1.114.$$

Figures 3 to 6 show the polar plots for $\boldsymbol{\mu}_\perp$ and $\boldsymbol{\mu}_A$, the evolution of the orthogonal isoclines and asymptotes and the geometry of the Newtonian asymptote.

The hodographs of $\mathbf{v}_t$ for a set of representative values of μ and initial $\mathbf{v}_t$ are shown in Fig. 8.

10 Conclusions

A thorough study of the phase-space geometry of the sliding velocity $\mathbf{v}_t$ flow for single-point rough collisions in multibody systems has been presented. Its dependence upon friction coefficient μ and five parameters associated with the collision configuration has been shown, and bifurcation and other significative values of μ have been determined.

A plane of vectors $\boldsymbol{\mu} \equiv \mu\boldsymbol{\sigma}$, being $\boldsymbol{\sigma}$ the unit vector in the sliding direction, has been introduced that greatly helps to characterize the phase space geometry. The knowledge of this geometry gives a full picture of the collision behavior for the configuration considered and for arbitrary initial generalized velocities. It is also useful to design robust computer codes to deal with rough collisions in multibody systems.

Acknowledgments

The author expresses his thanks to Ana Barjau for her useful comments and suggestions on this work and to Salvador Cardona for his collaboration in the numerical and graphical work implied in the application example.

References

Batlle, J. A., 1996, "Rough Balanced Collisions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 63, pp. 168–172.

Batlle, J. A., 1993, "On Newton's and Poisson's Rules of Percussive Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 60, pp. 376–381.

Batlle, J. A., and Condomines, A. B., 1991, "Rough Collisions in Multibody Systems," *Mech Mach Theory,* Vol. 26, No. 6, pp. 565–577.

Beghin, H., 1951, *Cours de Mécanique Théorique et Appliquée,* Gauthier-Villars, Paris, pp. 477–783.

Bhatt, V., and Koechling, J., 1993, "Three dimensional frictional rigid body impact," Mechanical Engineering Department, Cornell University, MSD-93-01, Dec. 1993; ASME JOURNAL OF APPLIED MECHANICS, accepted for publication.

Smith, C. E., 1991, "Predicting Rebounds Using Rigid-Body Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 754–758.

Stronge, W. J., 1994, "Swerve during three-dimensional impact of rough rigid bodies," ASME JOURNAL OF APPLIED MECHANICS, Vol. 60, pp. 605–611.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

An Energy-Based Approach to Computing Resonant Nonlinear Normal Modes

M. E. King
Assistant Professor,
Department of Aerospace and Mechanical Engineering,
Boston University,
Boston, MA 02215
Assoc. Mem. ASME

A. F. Vakakis
Associate Professor,
Department of Mechanical and Industrial Engineering,
University of Illinois at Urbana-Champaign,
Urbana, IL 61801
Assoc. Mem. ASME

A formulation for computing resonant nonlinear normal modes (NNMs) is developed for discrete and continuous systems. In a canonical framework, internal resonance conditions are immediately recognized by identifying commensurable linearized natural frequencies of these systems. Additionally, a canonical formulation allows for a single (linearized modal) coordinate to parameterize all other coordinates during a resonant NNM response. Energy-based NNM methodologies are applied to a canonical set of equations and asymptotic solutions are sought. In order to account for the resonant modal interactions, it will be shown that high-order terms in the $O(1)$ solutions must be considered (in the absence of internal resonances, a linear expansion at $O(1)$ is sufficient). Two applications ('3:1' resonances in a two-degree-of-freedom system and '3:1' resonance in a hinged-clamped beam) are then considered by which to demonstrate the resonant NNM methodology. It is shown that for some responses, nonlinear modal relations do not exist in the context of physical coordinates and thus a transformation to a canonical framework is necessary in order to appropriately define NNM relations.

1 Introduction

Nonlinear normal modes (NNMs) were originally studied by Rosenberg (1966) as a direct extension of (real) normal modes of classical vibration theory. Rosenberg defined a normal mode motion in a discrete system as one in which (i) the motions of all coordinates of the system are periodic and of the same period, and (ii) all coordinates reach their respective maxima simultaneously (i.e., the system "vibrates in unison"). General existence theorems for NNMs in discrete systems with analytic Hamiltonians were derived by Rosenberg (1964), Cooke and Struble (1966), and Weinstein (1973). Analytical perturbation methodologies were developed by Manevitch and Mikhlin (1972) and Mikhlin (1974) to study the sets of nonlinear, singular functional equations describing nonsimilar NNMs. An alternative framework for computing NNMs was introduced by Shaw and Pierre (1991, 1993) by constructing analytic approximations to invariant manifolds of the normal mode motion. This definition allows for the notion of NNM to be applied to both undamped and damped discrete oscillators. Their formulation resembles the construction of center manifolds close to nonhyperbolic fixed points of discrete nonlinear systems, and can be effectively used for studying NNMs in systems with no low-order internal resonances. Bifurcations of nonlinear normal modes in discrete oscillators have been examined in the works of Rand (1971), Month and Rand (1977), Caughey et al. (1990), Vakakis (1990), and Rand et al. (1992).

Applications of NNMs to continuous nonlinear systems have recently been addressed in the literature. Bennouna and White (1984) and Benamar et al. (1991) obtained normal mode solutions by considering discretized equations of motion. In Shaw and Pierre (1992), a Galerkin methodology was first used to discretize the equation of the nonlinear continuum, and the nonlinear modes of the resulting set of ordinary differential equations were computed by the invariant manifold approach. Several NNM methodologies have also been proposed which are applied directly to the governing nonlinear partial differential equations. Shaw and Pierre (1994) extended the invariant manifold methodology to investigate NNMs of bounded, nonlinear continua. A comparison between the continuous formulation and the discretized formulation was made by Boivin, Pierre, and Shaw (1994). An energy-based NNM methodology was developed by King and Vakakis (1994) for a class of one-dimensional, conservative, continuous systems. By imposing conservation of energy, the displacement of arbitrary material points can be expressed as a function of the displacement of a single reference point. An extension of this methodology to coupled continuous systems was developed by King and Vakakis (1995). Additional approaches for computing NNMs in continuous systems based on the method of multiple scales and on a complex-variable invariant manifold construction were developed by Nayfeh and Nayfeh (1994), and equivalence among the methods was demonstrated. A general exposition on the applications of NNMs in discrete and continuous systems is presented in Vakakis et al. (1995).

The aforementioned NNM methodologies fail when applied to systems possessing internal resonances (i.e., systems with commensurable linearized natural frequencies). To overcome this restriction, Nayfeh and Chin (1994) recently developed a methodology to compute NNMs in systems with internal resonances based on a complex-variable invariant manifold approach. In that work, systems with weak cubic nonlinearities were considered for cases of '1:1' and '3:1' internal resonances, and bifurcations of the computed NNMs were investigated. Rand (1995) considered '3:1' internal resonances using the method of multiple scales. In the present work, a extension of the energy-based NNM methodology (King and Vakakis, 1994) is developed for computing internally resonant NNMs. The formulation is carried out in a transformed canonical coordinate system, and high-order expansions are introduced at $O(1)$ so as to allow for resonant interactions. Using the displacement of the lowest resonant (linearized) mode as a reference, all other modal displacements are expressed with respect to this reference coordinate through resonant modal functions. Following the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 10, 1995; final revision, Oct. 22, 1995. Associate Technical Editor: S. W. Shaw.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

development of the general methodology, resonant NNMs are computed for two systems which possess '3:1' resonance conditions. For both applications, stability characteristics of the computed resonant NNMs are also investigated.

2 General Methodology

Consider the following system of equations in canonical form:

$$\ddot{u}_j + \omega_j^2 u_j + \epsilon f_j(\mathbf{u}) = 0, \quad j = 1, \ldots, n \tag{1}$$

where $\mathbf{u} = (u_1(t), u_2(t), \ldots, u_n(t))^T$ denotes the vector of canonical (modal) coordinates, ϵ is a perturbation parameter ($|\epsilon| \ll 1$), ω_j denotes the jth linearized natural frequency, and $f_j(\mathbf{u})$ denotes the nonlinear term of the jth equation of motion. For the present analysis, we assume that $f_j(\mathbf{u})$: (i) is obtained from a potential energy, $V(\mathbf{u})$, such that $\omega_j^2 u_j + \epsilon f_j(\mathbf{u}) = \partial V / \partial u_j$ (i.e., the system is conservative), and (ii) possesses only odd-order nonlinearities. Note that the equations for a general (conservative) mechanical system can be expressed as (1) using an appropriate modal transformation. Once the resonant NNM solutions are computed, the modal transformation may again be used to determine the corresponding physical displacements.

Choosing the kth coordinate in (1) as the *reference coordinate,* the response of any arbitrary coordinate is parameterized during a resonant NNM by the following functional relation (Rosenberg, 1966):

$$u_m(t) = U_m(u_k(t)), \quad m = 1, \ldots, n, \tag{2}$$

where $U_m(\cdot)$ is the mth "resonant modal function." In writing (2), the independent variable is changed from time, t, to the reference displacement, u_k, and thus the explicit time dependence of the motion is eliminated. Substituting (2) into (1), one obtains the following equation governing the mth resonant modal function:

$$U_m'' \dot{u}_k^2 + U_m' \ddot{u}_k + \omega_m^2 U_m + \epsilon F_m(\mathbf{U}, u_k) = 0 \tag{3}$$

where $(\cdot)' = d(\cdot)/du_k$, $\mathbf{U} = (u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_n)$ and $F_m(\mathbf{U}, u_k) \equiv f_m(\mathbf{u})|_{u_j = U_j}$. Considering the kth (reference) equation of motion, reference velocities and accelerations may also be obtained:

$$\dot{u}_k^2 = -2 \int_{u_k^*}^{u_k} [\omega_k^2 \xi + \epsilon F_k(\mathbf{U}(\xi), \xi)] d\xi \tag{4a}$$

$$\ddot{u}_k = -\omega_k^2 u_k - \epsilon F_k(\mathbf{U}(u_k), u_k) \tag{4b}$$

where u_k^* denotes the maximum amplitude attained by the reference coordinate u_k. Note that equation (4a) represents the kinetic energy of the reference coordinate; this quantity is identically zero, and the potential energy is maximal, when the reference coordinate reaches its maximum amplitude, $u_k = u_k^*$. Substituting (4a, b) into (3), one obtains the following equation for the mth nonlinear resonant modal function:

$$\left\{ 2 \int_{u_k}^{u_k^*} [\omega_k^2 \xi + \epsilon F_k(\mathbf{U}(\xi), \xi)] d\xi \right\} U_m'' - \{ \omega_k^2 u_k + \epsilon F_k(\mathbf{U}(u_k), u_k) \} U_m' + \omega_m^2 U_m + \epsilon F_m(\mathbf{U}, u_k) = 0 \tag{5}$$

where $m = 1, 2, \ldots, k-1, k+1, \ldots, n$. Note that Eq. (5) becomes singular at the maximum potential energy level ($u_k = u_k^*$) since the coefficients of the highest derivative, U_m'', vanish there. In the following analysis, asymptotic approximations to the resonant modal functions are developed for small reference displacements (i.e., for potential energies well below the maximum level), and are then analytically continued up to the maximum potential energy level. A similar analysis was performed by Manevitch and Mikhlin (1972), Mikhlin (1974) and Vakakis (1992b) for discrete nonlinear oscillators and by King and Vakakis (1994) for a class of nonlinear continuous systems. The following condition is thus imposed at the maximum potential energy level:

$$-\{ \omega_k^2 u_k^* + \epsilon F_k(\mathbf{U}(u_k^*), u_k^*) \} U_m'(u_k^*) + \omega_m^2 U_m(u_k^*) + \epsilon F_m(\mathbf{U}(u_k^*), u_k^*) = 0. \tag{6}$$

Condition (6) guarantees the analytic continuation of the asymptotic solutions up to the maximum potential energy level.

Asymptotic solutions of (5) are sought in the form of power series in terms of the reference displacement, u_k. The coefficients of this expansion are then determined so that the condition of analytic continuation (6) is satisfied when $u_k = u_k^*$. Away from the maximum potential energy level, a solution of the following form is introduced:

$$U_m(u_k) = U_m^{(0)}(u_k) + \epsilon U_m^{(1)}(u_k) + O(\epsilon^2) \tag{7a}$$

where

$$U_m^{(j)}(u_k) = c_{m,1}^j u_k + c_{m,3}^j u_k^3 + \ldots = \sum_{i=1,3,5,\ldots} c_{m,i}^j u_k^i \tag{7b}$$

for $m = 1, \ldots, k-1, k+1, \ldots, n$. *In contrast to the nonresonant methodology (King and Vakakis, 1994), which assumes linearity at $O(1)$, the present analysis allows for an arbitrary $O(1)$ expansion so as to fully account for any $O(1)$ resonant interactions.* Only odd-order terms arise in (7b) due to the earlier assumption of only odd-powered nonlinearities. Due to compatibility between the kth resonant nonlinear modal function, U_k, and the reference displacement, u_k, the following *compatibility conditions* must also be satisfied: $U_k^{(0)}(u_k) = u_k$ and $U_k^{(j)}(u_k) = 0$, $j \geq 1$. Substituting (7a) into (5) and (6), one obtains a hierarchy of subproblems at each order of ϵ, each of which will be considered separately.

***O*(1) Approximation:** Substituting (7a) into (5) and considering only terms of $O(1)$, one obtains the following equation governing the $O(1)$ approximation to the mth modal function:

$$\omega_k^2 (u_k^{*2} - u_k^2) U_m^{(0)\prime\prime} - \omega_k^2 u_k U_m^{(0)\prime} + \omega_m^2 U_m^{(0)} = 0 \tag{8}$$

which is complemented by the following $O(1)$ maximum potential energy condition:

$$-\omega_k^2 u_k U_m^{(0)\prime} + \omega_m^2 U_m^{(0)}|_{u_k = u_k^*} = 0 \tag{9}$$

where $m = 1, \ldots, k-1, k+1, \ldots, n$. Substituting (7$b$) into (8) and considering like-powered terms of u_k, one may obtain the following recursion relation for $c_{m,j}^{(0)}$:

$$c_{m,j}^{(0)} = \left[\frac{(j-2)^2 \omega_k^2 - \omega_m^2}{j(j-1)\omega_k^2 u_k^{*2}} \right] c_{m,j-2}^{(0)}, \quad j = 3, 5, 7, \ldots. \tag{10}$$

Using (10), any arbitrary coefficient, $c_{m,j}^{(0)}$, in the $O(1)$ expansion (7b) may be expressed solely in terms of the linear coefficient, $c_{m,1}^{(0)}$, as

$$c_{m,j}^{(0)} = \left[\frac{c_{m,1}^{(0)}}{j! \omega_k^{j-1} u_k^{*\,j-1}} \right] \prod_{n=2,4,\ldots}^{j-1} [(j-n)^2 \omega_k^2 - \omega_m^2], \quad j = 3, 5, 7, \ldots. \tag{11}$$

The problem is thus reduced to computing the linear coefficients of the $O(1)$ expansion, $c_{m,1}^{(0)}$. This coefficient may be determined using the maximum potential energy condition (9). Introducing (11) into (9), one obtains the following condition of analytic continuation:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\left[(\omega_k^2 - \omega_m^2) + \sum_{j=3,5,7,\ldots} \left\{ \frac{j\omega_k^2 - \omega_m^2}{j!\omega_k^{j-1}} \times \prod_{n=2,4,\ldots}^{j-1} [(j-n)^2\omega_k^2 - \omega_m^2] \right\} \right] u_k^* c_{m,1}^{(0)} = 0. \quad (12)$$

Truncating expansion (7*b*) at the *N*th order, Eq. (12) is further reduced to the following simplified form:

$$\frac{u_k^* c_{m,1}^{(0)}}{N!\omega_k^{N-1}} \prod_{i=1,3,\ldots}^{N} (\omega_m^2 - i^2\omega_k^2) = 0. \quad (13)$$

Careful consideration of Eq. (13) provides considerable insight into the problems associated with computing NNMs in the presence of internal resonances. Neglecting the trivial solution $u_k^* = 0$, two distinct classes of solutions at $O(1)$ are noted.

• *No Resonance Between mth and kth Coordinates* ($\omega_m \neq i\omega_k$, $i = 1, 3, \ldots, m \neq k$)

(i) The only nontrivial solution to (13) is $c_{m,1}^{(0)} = 0$.

(ii) The solution for the *m*th nonlinear modal function is therefore given by

$$U_m(u_k) = 0 + O(\epsilon). \quad (14)$$

(iii) Thus, in the absence of an internal resonance condition, no $O(1)$ interactions occur between these two canonical coordinates, and the problem degenerates to that studied by King and Vakakis (1994).

• *'i:1' Resonance Between mth and kth Coordinates* ($\omega_m = i\omega_k$ for some $i = 1, 3, 5, \ldots$)

(i) Equation (13) is identically satisfied and no additional information can be obtained at this order concerning $c_{m,1}^{(0)}$.

(ii) Using (11) and (12), the solution for the *j*th coefficient of (7*b*) is given by

$$c_{m,j}^{(0)} = \begin{cases} c_{m,1}^{(0)}, & j = 1 \\ \left[\dfrac{c_{m,1}^{(0)}}{j!\omega_k^{j-1}u_k^{*j-1}}\right] \displaystyle\prod_{n=2,4,\ldots}^{j-1} [(j-n)^2\omega_k^2 - \omega_m^2], & j = 3, 5, \ldots, i \\ 0, & j \geq i+2. \end{cases} \quad (15)$$

Note that for an '*i*:1' resonance, the $O(1)$ series expansion need only contain terms up to the *i*th power in order to fully account for the resonant interactions.

(iii) Thus, the $O(1)$ solution for the resonant coordinate can only be fully determined by considering $O(\epsilon)$ terms.

Hence, two distinctly different responses are predicted depending on the absence or presence of internal resonance conditions. To fully determine the $O(1)$ solution for a *resonant* modal function (i.e., to compute $c_{m,1}^{(0)}$), an $O(\epsilon)$ analysis must be carried out. In the following discussion, the $O(\epsilon)$ approximation is developed for *n*-degree-of-freedom systems which possess only a single resonance condition. A few comments are then made concerning extensions for the case of multiple resonances.

$O(\epsilon)$ Approximation: Introducing (7*a*) into Eqs. (5) and (6), the $O(\epsilon)$ approximation to the *m*th nonlinear modal function is governed by the following equation:

$$\omega_k^2(u_k^{*2} - u_k^2)U_m^{(1)\prime\prime} + \left(2\int_{u_k}^{u_k^*} F_k(\mathbf{U}^{(0)}(\xi), \xi)d\xi\right)U_m^{(0)\prime\prime} - \omega_k^2 u_k U_m^{(1)\prime} - F_k(\mathbf{U}^{(0)}, u_k)U_m^{(0)\prime} + \omega_m^2 U_m^{(1)} + F_m(\mathbf{U}^{(0)}, u_k) = 0 \quad (16a)$$

and is complemented by the following $O(\epsilon)$ maximum potential energy relation:

$$-\omega_k^2 u_k U_m^{(1)\prime} - F_k(\mathbf{U}^{(0)}, u_k)U_m^{(0)\prime} + \omega_m^2 U_m^{(1)} + F_m(\mathbf{U}^{(0)}, u_k)|_{u_k=u_k^*} = 0 \quad (16b)$$

where $m = 1, \ldots, k-1, k+1, \ldots, n$.

• *Single Resonance Condition.* The methodology is now developed for systems possessing only two resonant modes (i.e., a single '*p*:1' resonance between the reference and the *q*th linearized modes, $\omega_q = p\omega_k$). For such a system, only two modes interact at $O(1)$ during a NNM response, and the $O(1)$ solution may be summarized as follows (cf. Eqs. (14) and (15)):

$$U_m^{(0)}(u_k) = \begin{cases} u_k, & m = k \\ \displaystyle\sum_{j=1,3,\ldots}^{p} c_{q,j}^{(0)}(c_{q,1}^{(0)})u_k^j, & m = q \\ 0, & m = 1, \ldots, n, \quad m \neq k, \quad m \neq q \end{cases} \quad (17)$$

where the dependence on $c_{q,1}^{(0)}$ is noted. Moreover, since only two modes contribute at $O(1)$, the nonlinearities in (16*a*, *b*) may be expressed in following general form:

$$F_m(\mathbf{U}^{(0)}, u_k) = \sum_{j=1,3,5,\ldots} \left(\sum_{r+s=j, r,s\geq 0} \alpha_{m,r,s} u_k^r U_q^{(0)}(u_k)^s \right) = \sum_{j=1,3,5,\ldots} \beta_{m,j}(c_{q,1}^{(0)})u_k^j \quad (18)$$

where $\alpha_{m,r,s}$ denotes known coefficients from the original equations of motion, whereas $\beta_{m,j}(c_{q,1}^{(0)})$ denotes (as yet unknown) coefficients depending on $c_{q,1}^{(0)}$. Substituting (18) into (16*a*, *b*), one obtains a single equation governing the *resonant* coordinate ($m = q$ in (16*a*)):

$$\omega_k^2(u_k^{*2} - u_k^2)U_q^{(1)\prime\prime} + \left(2\int_{u_k}^{u_k^*} \left[\sum_{j=1,3,\ldots} \beta_{k,j}\xi^j\right]d\xi\right) \times \left[\sum_{j=1,3,\ldots}^{p} j(j-1)c_{q,j}^{(0)}j_k^{j-2}\right] - \omega_k^2 u_k U_q^{(1)\prime} - \left(\sum_{j=1,3,\ldots} \beta_{k,j}u_k^j\right)\left[\sum_{j=1,3,\ldots}^{p} jc_{q,j}^{(0)}u_k^{j-1}\right] + \omega_q^2 U_q^{(1)} + \left(\sum_{j=1,3,\ldots} \beta_{q,j}u_k^j\right) = 0 \quad (19a)$$

complemented by its respective maximum potential energy condition ($m = q$ in (16*b*)):

$$-\omega_k^2 u_k^* U_q^{(1)\prime}(u_k^*) - \left(\sum_{j=1,3,\ldots} \beta_{k,j}u_k^{*j}\right)\left[\sum_{j=1,3,\ldots}^{p} jc_{q,j}^{(0)}u_k^{*j-1}\right] + \omega_q^2 U_q^{(1)}(u_k^*) + \left[\sum_{j=1,3,\ldots} \beta_{q,j}u_k^{*j}\right) = 0. \quad (19b)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

There also exist $(n\text{-}2)$ equations for the *nonresonant* coordinates ($m \neq q$ in (16a)):

$$\omega_k^2(u_k^{*2} - u_k^2)U_m^{(1)\prime\prime} - \omega_k^2 u_k U_m^{(1)\prime} + \omega_m^2 U_m^{(1)} + \Big(\sum_{j=1,3,\ldots} \beta_{m,j}(c_{q,1}^{(0)})u_k^j\Big) = 0 \quad (20a)$$

complemented by their respective maximum potential energy relations ($m \neq q$ in (16b)):

$$-\omega_k^2 u_k^* U_m^{(1)\prime}(u_k^*) + \omega_m^2 U_m^{(1)}(u_k^*) + \Big(\sum_{j=1,3,\ldots} \beta_{m,j}(c_{q,1}^{(0)})u_k^j\Big) = 0 \quad (20b)$$

where $m = 1, \ldots, k-1, k+1, \ldots, q-1, q+1, \ldots, n$. Note that the above $(n\text{-}1)$ Eqs. (19a), (20a) are uncoupled and, thus, the problem reduces to a set of $(n\text{-}1)$ equations (plus their respective maximum potential energy conditions) which may be solved independently. It should be emphasized, however, that the solutions to (19a) and (20a) depend on the unknown coefficient $c_{q,1}^{(0)}$. This coefficient can be determined only by imposing the maximum potential energy relation for the resonant mode (19b). The remaining steps of the solution process are now described.

We will first consider the nonresonant coordinates ($m \neq q$). Introducing (7b) into (20a) and solving for like powers of u_k, the following recursion relation is obtained:

$$c_{m,j}^{(1)}(c_{q,1}^{(0)}) = \frac{1}{j(j-1)\omega_k^2 u_k^{*2}}[\{(j-2)^2\omega_k^2 - \omega_m^2\}c_{m,j-2}^{(1)}(c_{q,1}^{(0)}) - \beta_{m,j-2}^{(1)}(c_{q,1}^{(0)})], \quad j = 3, 5, 7, \ldots \quad (21)$$

where $m = 1, 2, \ldots, k-1, k+1, \ldots, q-1, q+1, \ldots, n$. Imposing the maximum potential energy relations (20b), one obtains the following conditions for the analytic continuation of the series expansions:

$$c_{m,1}^{(1)}(c_{q,1}^{(0)}) = \frac{1}{(\omega_k^2 - \omega_m^2)u_k^{*2}}\{\beta_{m,1}^{(1)}(c_{q,1}^{(0)}) + \sum_{j=3,5,\ldots}[\beta_{m,j}^{(1)}(c_{q,1}^{(0)}) - (j\omega_k^2 - \omega_m^2)u_k^{*j}c_{m,j}^{(1)}(c_{q,1}^{(0)})]\}. \quad (22)$$

Equations (21) and (22) may now be solved simultaneously for each nonresonant coordinate to obtain expressions for the $O(\epsilon)$ expansion coefficients, $c_{m,j}^{(1)}(c_{q,1}^{(0)})$ (note again the dependence of these coefficients on $c_{q,1}^{(0)}$).

The second task requires one to compute the $O(\epsilon)$ modal function for the resonant coordinate, $U_q^{(1)}(\bullet)$, and to simultaneously determine the $O(1)$ linear coefficient, $c_{q,1}^{(0)}$. Introducing the appropriate expansion (7b) into (19a), like-powered terms in u_k are equated and relations among the series expansion coefficients are determined. Using the results from this first step, the maximum potential energy condition (19b) is imposed. As with the $O(1)$ approximation, imposing the maximum potential energy condition fails to provide information concerning the $O(\epsilon)$ coefficient, $c_{q,1}^{(1)}$, and thus a $O(\epsilon^2)$ analysis must be performed (the details of which are not considered here). *The maximum potential energy relation does, however, provide an algebraic equation governing the O(1) expansion coefficient,* $c_{q,1}^{(0)}$. Based on the solutions of this algebraic equation, the $O(1)$ approximation to the resonant modal function and the $O(\epsilon)$ approximations to the $(n\text{-}2)$ *nonresonant* modal functions (using the solutions of Eqs. (21)(22)) may be fully determined.

The above analytic methodology can be carried out up to any desired order of approximation. Before proceeding to several applications of the above resonant NNM formulation, a few comments are in order concerning systems possessing multiple resonances and alternative techniques for computing resonant NNM.

• *Multiple Resonance Conditions.* Though a general formulation is not presented herein, it should be emphasized that the presented methodology may be directly extended to systems possessing *multiple resonance conditions;* one must merely allow for each resonant coordinate to introduce $O(1)$ effects. To this end, the $O(1)$ approximation to a resonant modal function must contain powers of the reference displacement up to the order of its respective resonance (i.e., a coordinate which is in a 'p:1' resonance with the reference must contain terms up to order u_0^p at $O(1)$). The remaining nonresonant coordinates will not contribute at $O(1)$. The solution at $O(\epsilon)$ is then merely a direct extension of the solution method discussed above.

An interesting observation should be noted. As with the nonresonant methodology (King and Vakakis, 1994), a single reference coordinate is used throughout the analysis to parameterize the resonant NNMs. Thus, the resulting motions may be conceived as occurring on a one-dimensional manifold in the system's configuration space, irrespective of the absence or presence of internal resonances. Though a resonant system was not considered in (Shaw and Pierre, 1993), it was noted in that paper that one must consider a $2N$ dimensional modal subspace (using the invariant-manifold approach) in order to capture the dynamics of N (resonantly) interacting modes. *As shown here, however, a one-dimensional manifold in configuration space is sufficient for resonant conservative systems.* It is therefore conjectured that the invariant manifold approach of (Shaw and Pierre, 1993) can be appropriately modified for conservative systems so as to compute two-dimensional invariant manifolds in phase space on which resonant normal mode motions exist, irrespective of the number of interacting modes. An implementation of this reduction in dimensionality could ease some of the computational issues associated with the invariant manifold approach in resonant systems.

Two applications are now considered in which to illustrate the resonant NNM methodology. In the first problem, '3:1' resonances are considered in a two-degree-of-freedom system with a single nonlinear grounding stiffness. The second problem considers '3:1' resonant interactions between the first and second linearized modes of a nonlinear hinged-clamped beam. For both applications, resonant NNMs are computed and stability characteristics of the resulting modes are investigated. Some of the computed resonant responses correspond to *strongly nonsimilar* NNMs, in contrast to previous works which only demonstrated *weakly nonsimilar* NNM behavior.

3 Application 1: '3:1' Resonance in a Two-Degree-of-Freedom System

Consider the following simple discrete system (Shaw and Pierre, 1993):

$$\ddot{x} + x + k(x - y) + \epsilon g x^3 = 0$$
$$\ddot{y} + y + k(y - x) = 0 \quad (23)$$

which possesses internal resonances when $k = 0$ ('1:1' resonance) or $k = 4$ ('3:1' resonance). In the present analysis, we will only consider the '3:1' resonance and will therefore introduce a detuning parameter, σ, through the relation: $k = 4 + \epsilon\sigma/2$ (refer to King (1995) for a discussion of the '1:1' resonance). To place system (23) in canonical form, the following (modal) coordinates are introduced:

$$u_1(t) = \frac{x(t) + y(t)}{2}, \quad u_2(t) = \frac{x(t) - y(t)}{2}. \quad (24)$$

Upon substituting (24) into (23), one obtains the following set of canonical equations:

$$\ddot{u}_1 + u_1 + \frac{\epsilon g}{2}(u_1 + u_2)^3 = 0$$
$$\ddot{u}_2 + (9 + \epsilon\sigma)u_2 + \frac{\epsilon g}{2}(u_1 + u_2)^3 = 0. \quad (25)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The resonant NNM formulation developed above is now applied to system (25).

Using $u_1(t)$ as the reference displacement, introducing the relation $u_2(t) = U_2(u_1(t))$, and following the general methodology, Eq. (25) becomes

$$\left\{2\int_{u_1}^{u_1^*}\left[\xi + \frac{\epsilon g}{2}(\xi + U_2(\xi))^3\right]d\xi\right\}U_2'' - \left\{u_1 + \frac{\epsilon g}{2}(u_1 + U_2(u_1))^3\right\}U_2' + (9 + \epsilon\sigma)U_2 + \frac{\epsilon g}{2}(u_1 + U_2)^3 = 0 \quad (26a)$$

and is complemented by the following maximum potential energy relation (at $u_1 = u_1^*$):

$$-\left\{u_1^* + \frac{\epsilon g}{2}(u_1^* + U_2(u_1^*))^3\right\}U_2'(u_1^*) + (9 + \epsilon\sigma)U_2(u_1^*) + \frac{\epsilon g}{2}(u_1^* + U_2(u_1^*))^3 = 0. \quad (26b)$$

Considering terms of $O(\epsilon)$ in (26a, b), the following equations for $U_2^{(1)}(u_1)$ are obtained:

$$(u_1^{*2} - u_1^2)U_2^{(1)\prime\prime} + g\left(\int_{u_1}^{u_1^*}(\xi + U_2^{(0)}(\xi))^3 d\xi\right)U_2^{(0)\prime\prime} - u_1 U_2^{(1)\prime} - \frac{g}{2}(u_1 + U_2^{(0)})^3 U_2^{(0)\prime} + 9U_2^{(1)} + \sigma U_2^{(0)} + \frac{g}{2}(u_1 + U_2^{(0)})^3 = 0 \quad (29a)$$

complemented by the $O(\epsilon)$ maximum energy relation:

$$-u_1^* U_2^{(1)\prime} - \frac{g}{2}(u_1 + U_2^{(0)})^3 U_2^{(0)\prime} + 9U_2^{(1)} + \sigma U_2^{(0)} + \frac{g}{2}(u_1 + U_2^{(0)})^3|_{u_1 = u_1^*} = 0. \quad (29b)$$

Utilizing (7b), symbolically equating like-powered terms of u_1 in (29a) via *MATHEMATICA*, and imposing the condition at maximum potential energy (29b), the following coefficients for the expansion of $U_2^{(1)}$ are computed (cf. Eq. (7b)):

$$c_{2,3}^{(1)} = \left(\frac{g}{3} - \frac{\sigma}{6u_1^{*2}}\right)c_{2,1}^{(0)} + \frac{g}{9}c_{2,1}^{(0)2} + \frac{g}{9}c_{2,1}^{(0)3} + \frac{7g}{405}c_{2,1}^{(0)4} - \frac{4}{3u_1^{*2}}c_{2,1}^{(1)}$$

$$c_{2,5}^{(1)} = -\frac{g\{1 - [(8\sigma/3gu_1^{*2}) - 2]c_{2,1}^{(0)} - 2c_{2,1}^{(0)3} - c_{2,1}^{(0)4}}{40u_1^{*2}}$$

$$c_{2,7}^{(1)} = -\frac{g\{1 - [(8\sigma/3gu_1^{*2}) - 7]c_{2,1}^{(0)} + 25c_{2,1}^{(0)2} + 33c_{2,1}^{(0)3} + 14c_{2,1}^{(0)4}}{105u_1^{*4}}$$

$$c_{2,9}^{(1)} = -\frac{g\{1 - [(8\sigma/3gu_1^{*2}) - 7]c_{2,1}^{(0)} - 3c_{2,1}^{(0)2} - 37c_{2,1}^{(0)3} - 28c_{2,1}^{(0)4}}{189u_1^{*6}}$$

$$c_{2,11}^{(1)} = -\frac{8gc_{2,1}^{(0)4}}{135u_1^{*8}} \quad \text{and} \quad c_{2,r}^{(1)} = 0 \quad \text{for} \quad r \geq 13 \quad (30)$$

Substituting (7a) into (26a, b) and considering only terms of $O(1)$, the following equations governing the $O(1)$ resonant modal function are obtained:

$$(u_1^{*2} - u_1^2)U_2^{(0)\prime\prime} - u_1 U_2^{(0)\prime} + 9U_2^{(0)} = 0 \quad (27a)$$

$$-u_1^* U_2^{(0)\prime}(u_1^*) + 9U_2^{(0)}(u_1^*) = 0. \quad (27b)$$

Introducing expansion (7b) into (27a), and using relation (15) for $i = 3$, one obtains the following results for the coefficients of the $O(1)$ expansion:

$$c_{2,3}^{(0)} = -\left(\frac{4c_{2,1}^{(0)}}{3u_1^{*2}}\right) \quad \text{and} \quad c_{2,j}^{(0)} = 0 \quad \text{for} \quad j \geq 5 \quad (28a)$$

which directly lead to the following solution for $U_2^{(0)}(u_1)$:

$$U_2^{(0)}(u_1) = c_{2,1}^{(0)}\left[u_1 - \left(\frac{4}{3u_1^{*2}}\right)u_1^3\right]. \quad (28b)$$

Again, $c_{2,1}^{(0)}$ is indeterminate at this order.

along with the following algebraic equation governing $c_{2,1}^{(0)}$:

$$c_{2,1}^{(0)3} - \frac{27}{17}c_{2,1}^{(0)2} - \left(\frac{24\sigma}{17gu_1^{*2}} - \frac{63}{17}\right)c_{2,1}^{(0)} + \frac{9}{17} = 0. \quad (31)$$

The roots of Eq. (31) are depicted in Fig. 1 for small ranges of the ratio σ/gu_1^{*2}. Since the canonical coordinates represent the (linearized) modal displacements for system (23), one observes that the branches labeled as '1' (for large positive mistunings) and '3' in Fig. 1 represent motions which possess significant interactions between the two linearized modes. In contrast, the branches labeled as '1' (for large negative mistunings) and '2' represent motions which are dominated by the first linearized mode, and thus the energy is confined to the reference mode. To complete the analysis to $O(\epsilon)$ (i.e., to compute $c_{2,1}^{(1)}$), one must now consider terms of $O(\epsilon^2)$. Using *MATHEMATICA*, the following relation for $c_{2,1}^{(1)}$ was obtained:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

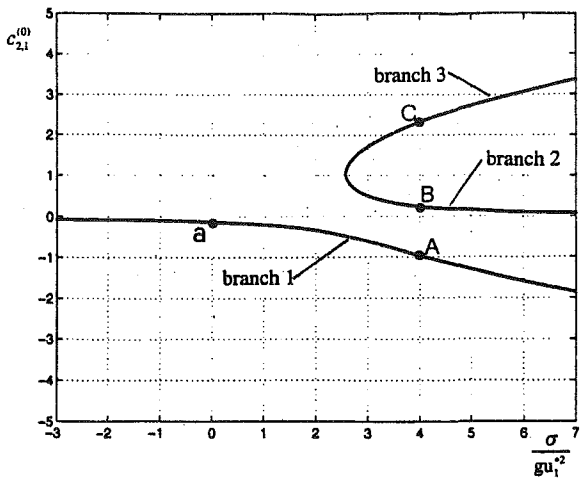


Fig. 1 Solutions for the linear coefficient, $c_{2,1}^{(0)}$, in the $O(1)$ approximation to the '3:1' resonant normal mode

$$\frac{-73\sigma}{4080} + \frac{233 g u_1^{*2}}{5440} + \left(\frac{89\sigma}{2040} - \frac{\sigma^2}{90 g u_1^{*2}} - \frac{31141 g u_1^{*2}}{152320}\right) c_{2,1}^{(0)}$$

$$+ \left(\frac{837827 g u_1^{*2}}{822528} - \frac{18811\sigma}{57120}\right) c_{2,1}^{(0)2}$$

$$+ \left(\frac{4006\sigma}{48195} - \frac{81077 g u_1^{*2}}{120960}\right) c_{2,1}^{(0)3}$$

$$+ \left(\frac{44401 g u_1^{*2}}{97920} - \frac{11633\sigma}{514080}\right) c_{2,1}^{(0)4}$$

$$- \frac{2497 g u_1^{*2}}{30464} c_{2,1}^{(0)5} + \frac{547 g u_1^{*2}}{26880} c_{2,1}^{(0)6}$$

$$+ \left[\frac{21}{17} - \frac{8\sigma}{17 g u_1^{*2}} - \frac{18 c_{2,1}^{(0)}}{17} + c_{2,1}^{(0)2}\right] c_{2,1}^{(1)} = 0 \quad (32)$$

i.e., a linear equation in $c_{2,1}^{(1)}$. Thus, one may easily compute $c_{2,1}^{(1)}$ for a given value of $c_{2,1}^{(0)}$ (obtained by solving Eq. (31)).

Combining all the above results, the approximation to the '3:1' resonant modal function, $U_2(u_1)$, is given by

$$U_2(u_1) = (c_{2,1}^{(0)} + \epsilon c_{2,1}^{(1)}) u_1(t) + (c_{2,3}^{(0)} + \epsilon c_{2,3}^{(1)}) u_1^3(t) + \epsilon c_{2,5}^{(1)} u_1^5(t) + \epsilon c_{2,7}^{(1)} u_1^7(t) + \epsilon c_{2,9}^{(1)} u_1^9(t) + \epsilon c_{2,11}^{(1)} u_1^{11}(t) + O(\epsilon^2) \quad (33)$$

where $c_{2,1}^{(0)}$ and $c_{2,1}^{(1)}$ are computed from Eq. (31) and (32), respectively, and $c_{2,j}^{(1)} (j \geq 3)$, are given by Eqs. (30). The three normal mode solutions corresponding to $\sigma = 1.0$, $g = 1.0$ and $u_1^* = 0.5$ (labeled as 'A', 'B' and 'C' in Fig. 1) are provided in Figs. 2(*a*), 3(*a*), and 4(*a*), respectively. The solutions depicted in figures 2–4(*a*) correspond to resonant NNMs in terms of canonical coordinates and thus provide some indication of the interactions between the resonant interactions of the system's linearized modes. From Figs. 2(*a*) and 4(*a*), we observe that the higher mode contributes significantly to the total response. Note that the computed resonant NNMs can be classified as *strongly nonsimilar* due to the highly nonlinear relationship required at $O(1)$ to capture the resonance effects.

Following a procedure similar to (King and Vakakis, 1994), the parameterizing variable is computed as

$$u_1(t) = u_1^* \cos\left[\{1 + \epsilon(\alpha_1 u_1^{*2} + \alpha_2 u_1^{*4} + \alpha_3 u_1^{*6} + \alpha_4 u_1^{*8}) + O(\epsilon^2)\} t\right] + O(\epsilon^2) \quad (34a)$$

where

$$\alpha_1 = \frac{3g}{16}(1 + c_{2,1}^{(0)})^3, \; \alpha_2 = \frac{15g}{32}(1 + c_{2,1}^{(0)})^2 c_{2,3}^{(0)},$$

$$\alpha_3 = \frac{105 g}{256}(1 + c_{2,1}^{(0)}) c_{2,3}^{(0)2} \quad \text{and} \quad \alpha_4 = \frac{63 g}{512} c_{2,3}^{(0)3}. \quad (34b)$$

Transforming (33) into physical coordinates, one obtains the following '3:1' resonant NNMs for system (23):

$$x(t) = u_1(t) + u_2(t) = (1 + c_{2,1}^{(0)} + \epsilon c_{2,1}^{(1)}) u_1^* \cos(\omega_n t) + (c_{2,3}^{(0)} + \epsilon c_{2,3}^{(1)}) u_1^{*3} \cos^3(\omega_n t) + \sum_{j=5,7,9,11} \epsilon c_{2,j}^{(1)} u_1^{*j} \cos^j(\omega_n t)$$

$$y(t) = u_1(t) - u_2(t) = (1 - c_{2,1}^{(0)} - \epsilon c_{2,1}^{(1)}) u_1^* \cos(\omega_n t) - (c_{2,3}^{(0)} + \epsilon c_{2,3}^{(1)}) u_1^{*3} \cos^3(\omega_n t) - \sum_{j=5,7,9,11} \epsilon c_{2,j}^{(1)} u_1^{*j} \cos^j(\omega_n t) \quad (35)$$

where $\omega_n = 1 + \epsilon(\alpha_1 u_1^{*2} + \alpha_2 u_1^{*4} + \alpha_3 u_1^{*6} + \alpha_4 u_1^{*8}) + O(\epsilon^2)$. The physical displacements during the resonant normal modes computed from (35) are depicted in Figs. 2–4(*b*) for the points labeled as 'A', 'B', and 'C', respectively, in Fig. 1. It is noted that in *physical coordinates, folding of the modal lines may result in multivalued relationships among the various coordi-*

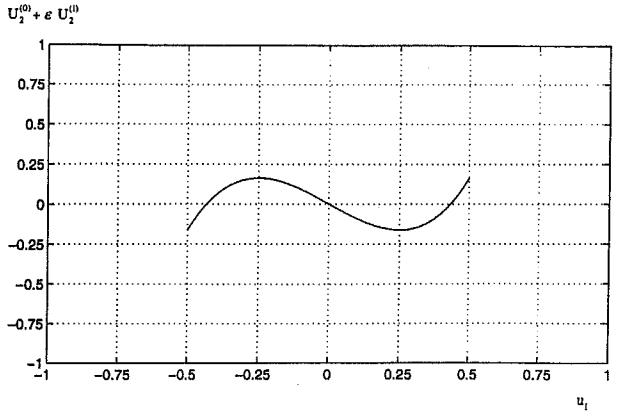


Fig. 2(*a*)

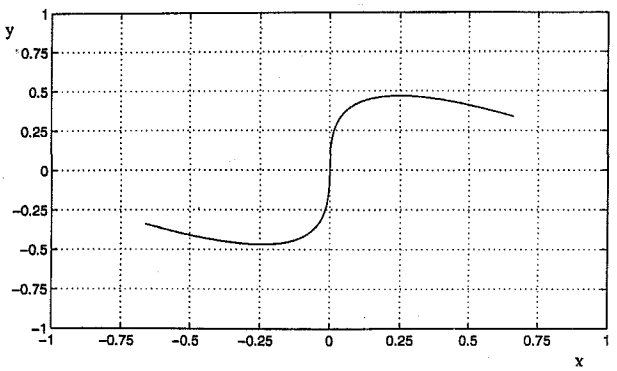


Fig. 2(*b*)

Fig. 2 Resonant NNM corresponding to point 'A' of Fig. 1; (*a*) canonical coordinates, and (*b*) physical coordinates

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

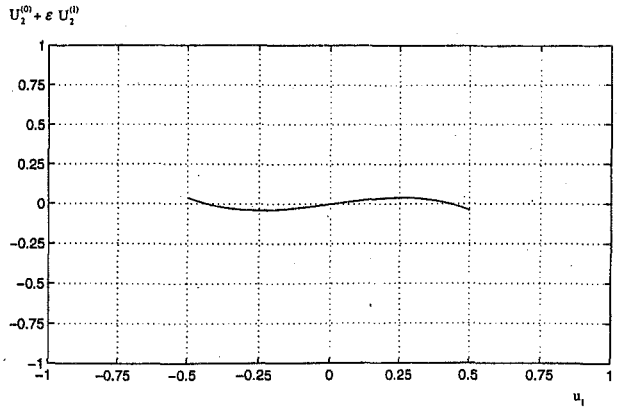


Fig. 3(*a*)

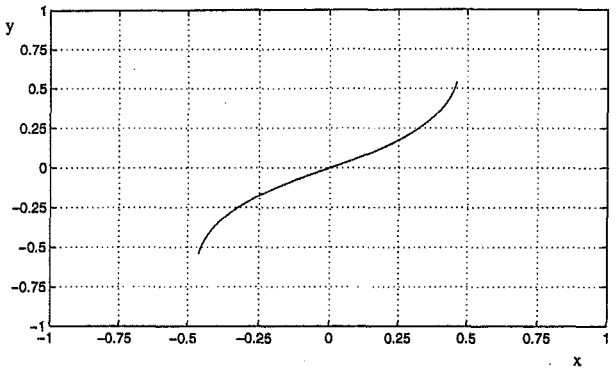


Fig. 3(*b*)

Fig. 3 Resonant NNM corresponding to point 'B' of Fig. 1; (*a*) canonical coordinates, and (*b*) physical coordinates

nates (cf. Fig. 4(*b*)). In such cases, a NNM-type expansion in terms of the physical coordinates breaks down. Note, also, that the modal lines in physical space fail to satisfy the definition of NNMs since (i) not all coordinates reach their maximum values simultaneously and (ii) a functional relation of the form (2) may not, in some cases, be valid for all times (due to multivaluedness). These two observations demonstrate the shortcomings of the general, nonresonant NNM formulation in the presence of internal resonances.

Stability characteristics of the computed resonant NNMs were discussed in King (1995). In summary, Poincare maps were constructed for various values of the mistuning parameter, and it was determined that the resonant NNM responses corresponding to branches 2 and 3 (of Fig. 1) are orbitally stable, and thus physically realizable. In contrast, motions corresponding to branch 1 (of Fig. 1) are only orbitally stable for mistunings below a critical mistuning value, given by $(\sigma/gu_1^{*2})_{\text{crit}} \approx 3.14$. Above this critical value, resonant modal responses denoted by branch 1 (in Fig. 1) are orbitally unstable, and thus *not* physically realizable. Typically, one would expect branches 1 and 3 to share the same stability characteristics, while branch 2 is of the opposite stability. In obtaining these results, however, one must recall that even though energy is conserved for a given solution, the energy levels associated with any two distinct solutions are in general not the same. Though the details are not presented here, one may obtain the expected saddle-node structure (and corresponding exchange of stability) by imposing a constant energy level for all computed solutions.

4 Application 2: '3:1' Resonance in a Hinged-Clamped Beam

A second application is now considered for a hinged-clamped nonlinear Euler-Bernoulli beam which possesses a '3:1' internal resonance between its first two linearized modes (Nayfeh and Mook, 1979). In nondimensional form, this system is governed by

$$w_{\tau\tau} + w_{ssss} = \epsilon\alpha w_{ss} \int_0^1 w_s^2 ds$$

$$w(0, \tau) = w_{ss}(0, \tau) = w(1, \tau) = w_s(1, \tau) = 0. \quad (36)$$

Equation (36) is discretized using the system's normalized linearized eigenfunctions, ϕ_k, by introducing the following expansion:

$$w(s, \tau) = \sum_{k=1,2,\ldots} \phi_k(s) u_k(\tau) \quad (37)$$

where $\phi_k = c_k[(\sinh \beta_k)(\sin \beta_k s) - (\sin \beta_k)(\sinh \beta_k s)$ for $k = 1, 2, 3, \ldots$ and β_k and c_k are chosen such that $\tan(\beta_k) - \tanh(\beta_k) = 0$ and $\int_0^1 \phi_k^2 ds = 1$. Substituting (37) into (36) and utilizing orthogonality of the eigenfunctions, one obtains the following set of discretized equations:

$$\ddot{u}_i + \omega_i^2 u_i = \epsilon\alpha \sum_j \sum_k \sum_l \Gamma_{ijkl} u_j u_k u_l \quad (38)$$

where u_i and ω_i $(\equiv \beta_i^2)$ denote the ith (linearized) modal displacement and natural frequency, respectively, and $\Gamma_{ijkl} \equiv \int_0^1 \phi_i(s)\phi_j''(s)ds \int_0^1 \phi_k'(s)\phi_l'(s)ds$. The first two linearized natural frequencies are nearly commensurable since $\omega_2 = (3 + \epsilon\sigma)\omega_1$, where $\epsilon\sigma = 0.241$. Thus, a '3:1' internal resonance condition exists. Parameterizing arbitrary modal displacements with respect to the first mode $(u_i(\tau) = U_i(u_1(\tau)))$, one obtains

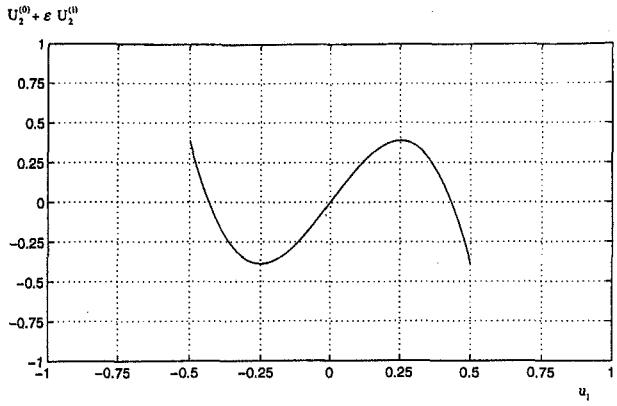


Fig. 4(*a*)

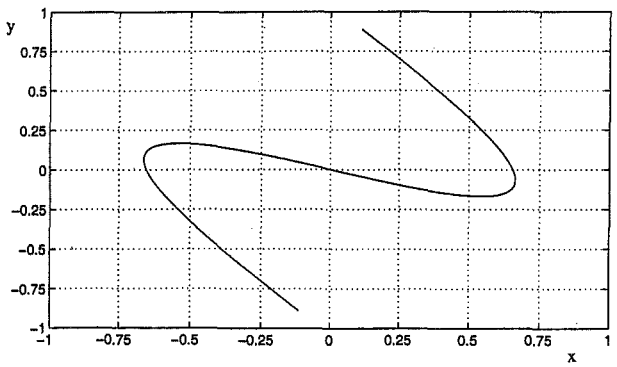


Fig. 4(*b*)

Fig. 4 Resonant NNM corresponding to point 'C' of Fig. 1; (*a*) canonical coordinates, and (*b*) physical coordinates

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

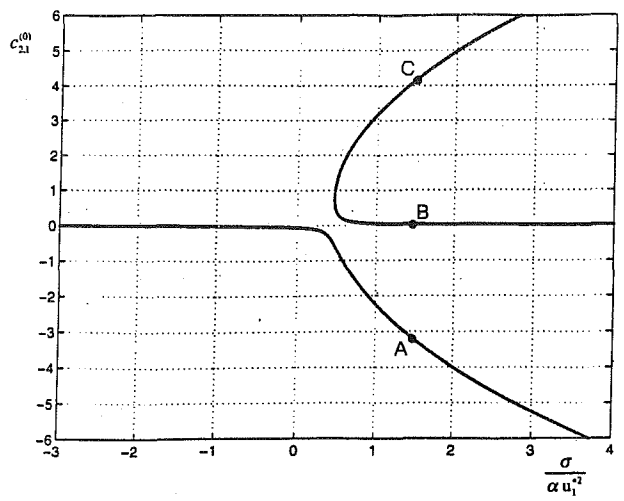


Fig. 5 Solutions for the linear coefficient, $c_{2,1}^{(0)}$, in the $O(1)$ approximation to the '3:1' resonant modal function for the hinged-clamped beam

$$U_i''\left\{\omega_1^2(u_1^{*2}-u_1^2)-2\epsilon\alpha\sum_{j,k,l}\Gamma_{1jkl}\int_{u_1}^{u_1^*}U_j(\xi)U_k(\xi)U_l(\xi)d\xi\right\}$$

$$-U_i'\left\{\omega_1^2u_1-\epsilon\alpha\sum_{j,k,l}\Gamma_{1jkl}U_jU_kU_l\right\}+\omega_i^2U_i$$

$$=\epsilon\alpha\sum_{j,k,l}\Gamma_{ijkl}U_jU_kU_l \quad (39)$$

where $i = 1, 2, 3, \ldots$, complemented by a set of maximum potential energy conditions (cf. Eq. (6)). Substituting the asymptotic expansions (7*a*) into (39), one obtains a hierarchy of subproblems at each order of ϵ.

Since $\omega_j \neq p\omega_1$, for $j \geq 3$ and p a small integer (i.e., no other low-order resonances exist), one may directly employ Eqs. (7*b*) and (15) to obtain

$$U_i^{(0)}(u_1)=\begin{cases}u_1, & \text{for } i=1\\ c_{2,1}^{(0)}\left(u_1-\dfrac{4}{3u_1^{*2}}u_1^3\right), & \text{for } i=2\\ 0, & \text{for } i\geq 3\end{cases} \quad (40)$$

where $c_{2,1}^{(0)}$ is indeterminate at this order of approximation. General solutions for the $O(\epsilon)$ approximations to the resonant modal functions are now sought in the following form:

$$U_i^{(1)}(u_1)=\begin{cases}0, & \text{for } i=1\\ \displaystyle\sum_{j=1,3,\ldots}c_{i,j}^{(1)}u_1^j, & \text{for } i\neq 1.\end{cases} \quad (41)$$

Using *MATHEMATICA,* it was determined that terms only up to the 11th-order in u_1 are required in (41) (i.e., $c_{i,j}^{(1)} = 0$ for $j \geq 13$ and $i \neq 1$). Since the resulting set of equations at $O(\epsilon)$ are uncoupled, general solutions for any arbitrary mode may be obtained (i.e., no truncation of the modal expansion (37) is required at this order). The system of equations resulting from the substitution of (41) into the $O(\epsilon)$ equations were solved using *MATHEMATICA* for the lowest six linearized modes (see King (1995) for further details on this analysis). Based on these results, the following equation was obtained for $c_{2,1}^{(0)}$:

$$c_{2,1}^{(0)3}-0.99111c_{2,1}^{(0)2}+\left(5.6235-12.744\frac{\sigma}{\alpha u_1^{*2}}\right)c_{2,1}^{(0)}+0.33037=0. \quad (42)$$

As in the previous application, one obtains a cubic polynomial for $c_{2,1}^{(0)}$. Solutions to (42) are depicted in Fig. 5; note that one obtains either one or three solutions, depending on the particular value of the detuning parameter.

It should again be emphasized that the results obtained at $O(\epsilon)$ are incomplete due to the indeterminacy $c_{2,1}^{(1)}$. Using *MATHEMATICA,* $c_{2,1}^{(1)}$ was computed by considering terms at $O(\epsilon^2)$. Though the details of this analysis are not presented here, it should be noted that during this solution process, a truncation of the infinite modal expansion (37) was necessary. A convergence study was performed for various modal truncations (with $\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$) and sufficiently accurate solutions were obtained by considering only the lowest nine modes. A summary of the convergence of $c_{2,1}^{(1)}$ is presented in Table 1.

Combining the above results, the following approximations to the resonant modal functions were obtained:

$$U_2(u_1)=(c_{2,1}^{(0)}+\epsilon c_{2,1}^{(1)})u_1+(c_{2,3}^{(0)}+\epsilon c_{2,3}^{(1)})u_1^3+\epsilon c_{2,5}^{(1)}u_1^5+\epsilon c_{2,7}^{(1)}u_1^7+\epsilon c_{2,9}^{(1)}u_1^9+\epsilon c_{2,11}^{(1)}u_1^{11}+O(\epsilon^2) \quad (43a)$$

$$U_i(u_1)=\epsilon c_{i,1}^{(1)}u_1+\epsilon c_{i,3}^{(1)}u_1^3+\epsilon c_{i,5}^{(1)}u_1^5+\epsilon c_{i,7}^{(1)}u_1^7+\epsilon c_{i,9}^{(1)}u_1^9+\epsilon c_{i,11}^{(1)}u_1^{11}+O(\epsilon^2),\ i\geq 3. \quad (43b)$$

Numerical values for the coefficients in (43*a*, *b*) were computed for $\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$, and the resulting lowest three modal functions (one resonant and two nonresonant) are graphically depicted in Figs. 6 through 8. Figure 6 corresponds to a typical NNM on branch 1 of Fig. 5 (labeled as 'A'). Similarly, Figs. 7 and 8 correspond to points 'B' and 'C', respectively, as labeled in Fig. 5. Considering these figures, we note that the response of the nonresonant modal functions are minimal in comparison with the resonant mode.

Due to the high dimensionality of the hinged-clamped beam, stability characteristics of the computed resonant NNMs were determined by a method based on Floquet analysis (King and Vakakis, 1995). Denoting by $\tilde{w}(s, \tau)$ the motion of the beam during a resonant NNM, small perturbations are introduced as follows:

$$w(s,\tau)=\tilde{w}(s,\tau)+\epsilon\xi(s,\tau). \quad (44)$$

Substituting (44) into (36) and retaining only terms of $O(\epsilon)$, the following linear variational equation with time-periodic coefficients is obtained:

$$\xi_{\tau\tau}+\xi_{ssss}=\epsilon\alpha\xi_{ss}\int_0^1\tilde{w}_s^2(\zeta,\tau)d\zeta+2\epsilon\alpha\tilde{w}_{ss}\int_0^1\tilde{w}_\zeta(\zeta,\tau)\xi_\zeta(\zeta,\tau)d\zeta+O(\epsilon^2). \quad (45)$$

Table 1 Convergence results for $c_{2,1}^{(1)}$ based on various modal truncations

Order of modal truncation	$c_{2,1}^{(1)}$ corresponding to Point 'A'	$c_{2,1}^{(1)}$ corresponding to Point 'B'	$c_{2,1}^{(1)}$ corresponding to Point 'C'
6	-5.57438	-0.0505513	-5.67480
7	-5.57373	-0.0505472	-5.67443
8	-5.57340	-0.0505452	-5.67426
9	-5.57323	-0.0505441	-5.67417
10	-5.57313	-0.0505434	-5.67412
11	-5.57307	-0.0505431	-5.67410
12	-5.57303	-0.0505428	-5.67408

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

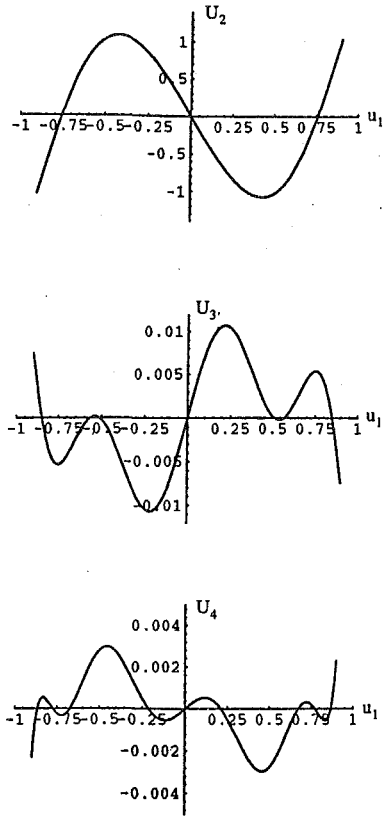


Fig. 6 Numerical solutions for the three lowest modal functions, corresponding to point 'A' of Fig. 5 ($\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$)

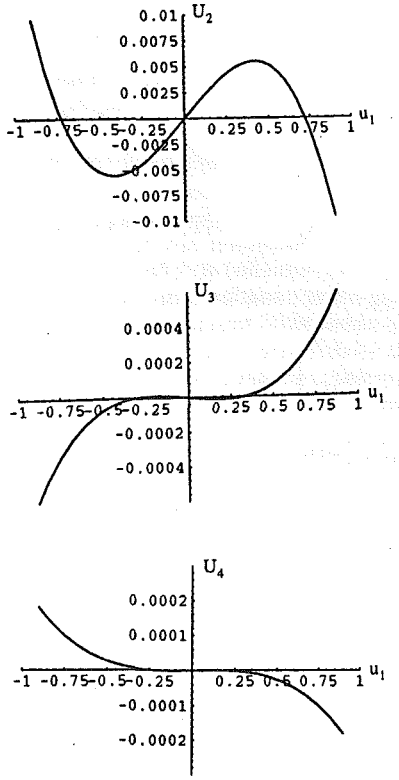


Fig. 7 Numerical solutions for the three lowest modal functions, corresponding to point 'B' of Fig. 5 ($\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$)

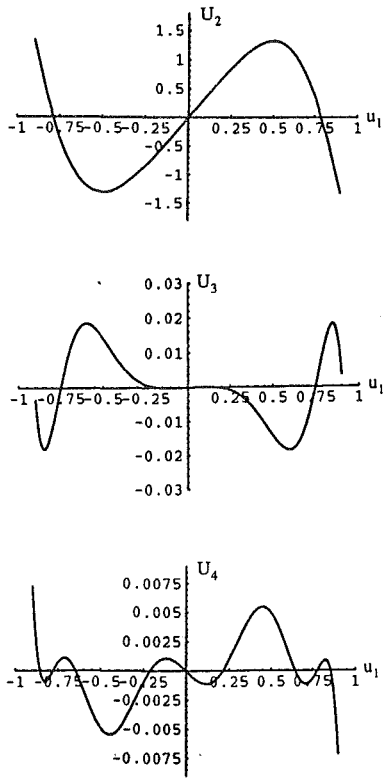


Fig. 8 Numerical solutions for the three lowest modal functions, corresponding to point 'C' of Fig. 5 ($\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$)

Expanding the solutions of (45) in terms of the system's linearized eigenfunctions ($\xi(s, \tau) = \sum_{k=1,2,\ldots} \phi_k(s) p_k(\tau)$), one may obtain the following infinite set of linear ordinary differential equations:

$$\ddot{p}_k + \omega_k^2 p_k = \epsilon\alpha\left(\int_0^1 \tilde{w}_s^2 ds\right)\left(\sum_{m=1,2,\ldots}\left[\int_0^1 \phi_k \phi_m'' ds\right] p_m\right) + 2\epsilon\alpha\left(\int_0^1 \phi_k \tilde{w}_{ss} ds\right)\left(\sum_{m=1,2,\ldots}\left[\int_0^1 \phi_m' \tilde{w}_s ds\right] p_m\right). \quad (46)$$

Floquet analysis (Nayfeh and Mook, 1979) is now performed by considering only the lowest N equations of (46) and computing the eigenvalues of the numerically computed Floquet matrix. Eigenvalues greater than unity in modulus indicate orbital instability, whereas eigenvalues of unity modulus indicate neutral stability. Stability or instability of the variational Eqs. (50) implies orbital stability or instability for the respective resonant NNM. Equation (43) was numerically integrated over one period and the resulting solution was utilized as the reference displacement. Using the following parameters: $\epsilon = 0.1$, $\alpha = 2.0$ and $u_1^* = 0.9$ (corresponding to the NNMs labeled as 'A', 'B', and 'C' in Fig. 5), the numerical Floquet analysis was applied and the resulting Floquet multipliers (considering only the lowest five terms in (46)) are provided in Table 2. Consider-

Table 2 Floquet multipliers resulting from the numerical stability analysis of the hinged-clamped beam (based on five mode truncation)

Multipliers associated with point 'A'	Multipliers associated with point 'B'	Multipliers associated with point 'C'
1.20917, 1.00000	1.00000, 1.00000	1.00000, 1.00000
1.00000, 1.00000	1.00000, 1.00000	1.00000, 1.00000
1.00000, 1.00000	1.00000, 1.00000	1.00000, 1.00000
1.00000, 1.00000	1.00000, 1.00000	1.00000, 1.00000
1.00000, **0.82701**	1.00000, 1.00000	1.00000, 1.00000

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ing Table 2, it is observed that point 'A' of Fig. 5 corresponds to an orbitally unstable resonant NNM due to the presence of a pair of Floquet multipliers which are not unity in modulus. Moreover, the points labeled as 'B' and 'C' in Fig. 5 are found to be orbitally stable, and thus physically realizable.

5 Discussion

An energy-based NNM methodology has been developed by which to study modal interactions introduced through internal resonances. The formulation is based on a discrete set of equations with uncoupled $O(1)$ linear parts (i.e., in modal, or canonical, form). Using a single canonical coordinate as a reference, the displacements of all other coordinates can be expressed through appropriately defined resonant modal functions. In contrast to the nonresonant NNM methodology (King and Vakakis, 1994), which validly assumes linearity for the $O(1)$ approximations to the modal functions, the presence of internal resonances requires high-order terms to be considered for the $O(1)$ approximations to the resonant modal functions. In fact, for a p:1 internal resonance condition, it has been shown that one must consider terms up to order p in the $O(1)$ expansion to fully account for the resonant effects.

The first application of the general theory considered a simple two-DOF system (3.1) possessing a '3:1' internal resonance. Depending on the system mistuning, either one or three coexisting resonant NNMs were computed. Comparing the resulting resonant modal lines in canonical and physical coordinates, the necessity of performing the analysis in the context of a canonical framework was demonstrated. It was noted that in physical configuration space, the modal lines do not satisfy the normal mode definition since not all coordinates reach their respective maxima simultaneously. The second application considered '3:1' resonant interactions in a nonlinear hinged-clamped beam. The results for the resonant modes of this continuous system were qualitatively similar to those obtained for the discrete system. Due to the high dimensionality of the phase space for this system, stability characteristics were ascertained for the computed resonant NNMs using Floquet analysis, and it was found that some resonant NNMs may be unstable.

Acknowledgments

This work was supported by an NSF Graduate Fellowship and by NSF Grant MSS 92-07318. Dr. Devendra Garg is the Grant monitor.

References

Benamar, R., Bennouna, M. M. K., and White, R. G., 1991, "The effects of large vibration amplitudes on the mode shapes and natural frequencies of thin elastic structures, Part I: Simply supported and clamped-clamped beams," *Journal of Sound and Vibration,* Vol. 149, pp. 179–195.

Bennouna, M. M. K., and White, R. G., 1984, "The effects of large vibration amplitudes on the fundamental mode shape of a clamped-clamped uniform beam," *Journal of Sound and Vibration,* Vol. 96, pp. 309–331.

Boivin, N., Pierre, C., and Shaw, S. W., 1993, "Nonlinear normal modes, invariance, and modal dynamics approximations of nonlinear systems," *Proceedings of the 14th Biennial Conference on Mechanical Vibration and Noise,* Albuquerque, NM, Sept. 19–22.

Caughey, T. K., Vakakis, A. F., and Sivo, J., 1990, "Analytical study of similar normal modes and their bifurcations in a class of strongly nonlinear systems," *International Journal of Non-Linear Mechanics,* Vol. 25, No. 5, pp. 521–533.

Cooke, C. H., and Struble, R. A., 1966, "On the existence of periodic solutions and normal mode vibrations of nonlinear systems," *Quarterly of Applied Mathematics,* Vol. 24, No. 3, pp. 177–193.

King, M. E., 1995, "Analytical and experimental aspects of nonlinear normal modes and nonlinear mode localization," Ph.D. Thesis, University of Illinois, Urbana, IL.

King, M. E., and Vakakis, A. F., 1994, "An Energy-Based Formulation for Computing Nonlinear Normal Modes in Undamped Continuous Systems," ASME *Journal of Vibration and Acoustics,* Vol. 116, pp. 332–340.

King, M. E., and Vakakis, A. F., 1995, "Asymptotic analysis of nonlinear mode localization in a class of coupled continuous structures," *Int. Journal of Solids and Structures,* Vol. 32, No. 8–9, pp. 1161–1177.

Manevitch, L. I., and Mikhlin, Y. V., 1972, "On periodic solutions close to rectilinear normal vibration modes," *PMM,* Vol. 36, No. 6, pp. 1051–1058.

Mikhlin, Y. V., 1974, "Resonance modes of near-conservative nonlinear systems," *PMM,* Vol. 38, No. 3, pp. 459–464.

Month, L. A., and Rand, R. H., 1977, "The Stability of Bifurcating Periodic Solutions in a Two Degree-of-Freedom Nonlinear System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, pp. 782–783.

Nayfeh, A. H., and Chin, C., 1994, "On Nonlinear Normal Modes of Systems With Internal Resonance," ASME *Journal of Vibration and Acoustics,* in press.

Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations,* John Wiley and Sons, New York.

Nayfeh, A. H., and Nayfeh, S. A., 1994, "On Nonlinear Modes of Continuous Systems," ASME *Journal of Vibration and Acoustics,* Vol. 116, pp. 129–136.

Rand, R. H., 1971, "A higher order approximation for Nonlinear Normal Modes in two-dof Systems," *International Journal of Non-Linear Mechanics,* Vol. 6, pp. 545–547.

Rand, R. H., 1995, personal communications with authors.

Rand, R. H., Pak, C. H., and Vakakis, A. F., 1992, "Bifurcations of nonlinear normal modes in a class of two degree of freedom systems," *Acta Mechanica,* [Suppl] Vol. 3, pp. 129–145.

Rosenberg, R. M., 1964, "On the existence of normal mode vibrations of nonlinear systems with two degrees of freedom," *Quarterly of Applied Mathematics,* Vol. 22, No. 3, pp. 217–234.

Rosenberg, R. M., 1966, "On non-linear vibrations of systems with many degrees of freedom," *Advances in Applied Mechanics,* Vol. 9, pp. 155–242.

Shaw, S. W., and Pierre, C., 1991, "Nonlinear normal modes and invariant manifolds," *Journal of Sound and Vibration,* Vol. 150, pp. 170–173.

Shaw, S. W., and Pierre, C., 1992, "On nonlinear normal modes," Winter Annual Meeting, ASME, Anaheim, CA, Nov. 8–13.

Shaw, S. W., and Pierre, C., 1993, "Normal modes for nonlinear vibratory systems," *Journal of Sound and Vibration,* Vol. 164, No. 1, pp. 85–96.

Shaw, S. W., and Pierre, C., 1994, "Normal modes of vibration for nonlinear continuous systems," *Journal of Sound and Vibration,* Vol. 169, No. 3, pp. 319–348.

Vakakis, A. F., 1990, "Analysis and identification of linear and nonlinear normal modes in vibrating systems," Ph.D. Thesis, California Institute of Technology, Pasadena, CA.

Vakakis, A. F., 1992a, "Fundamental and subharmonic resonances in a system with a '1-1' internal resonance," *Nonlinear Dynamics,* Vol. 3, pp. 123–143.

Vakakis, A. F., 1992b, "Nonsimilar normal oscillations in a strongly nonlinear discrete system," *Journal of Sound and Vibration,* Vol. 151, No. 22, pp. 341–361.

Vakakis, A. F., and Caughey, T. K., 1992, "A Theorem on the Exact Nonsimilar Steady-State Motions of a Nonlinear Oscillator," ASME JOURNAL OF APPLIED MECHANICS, Vol. 59, pp. 418–424.

Vakakis, A. F., Manevitch, L. I., Mikhlin, Y. V., Pilipchuk, V. N., and Zevin, A. A., 1995, *Normal Modes and Localization in Continuous Systems,* Wiley Interscience, in press.

Weinstein, A., 1973, "Normal modes of nonlinear hamiltonian systems," *Inventiones math.,* Vol. 20, pp. 47–57.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

H. Mang
Professor.

P. Helnwein
Assistant.

Institute for Strength of Materials,
Technical University of Vienna,
Vienna, Austria

R. H. Gallagher
Professor in Residence,
Aerospace and Mechanical Engineering
Department,
University of Arizona,
Tucson, AZ 85721

On the Nonuniqueness of Solutions Obtained With Simplified Variational Principles[1]

The attempt to satisfy subsidiary conditions in boundary value problems without additional independent unknowns in the form of Lagrange multipliers has led to the development of so-called "simplified variational principles." They are based on using the Euler-Lagrange equations for the Lagrange multipliers to express the multipliers in terms of the original variables. It is shown that the conversion of a variational principle with subsidiary conditions to such a simplified variational principle may lead to the loss of uniqueness of the solution of a boundary value problem. A particularly simple form of the geometrically nonlinear theory of bending of beams is used as the vehicle for this proof. The development given in this paper is entirely analytical. Hence, the deficiencies of the investigated simplified variational principle are fundamental.

1 Introduction

Frequently, the solution of boundary value problems requires consideration of subsidiary conditions. A simple way of accounting for such conditions in the corresponding variational principles is the Lagrange multiplier method (LMM). The multipliers represent a set of additional independent variables.

The attempt to satisfy subsidiary conditions without additional independent unknows has led to the development of so-called "simplified variational principles" (SVP). They are based on using the Euler-Lagrange equations for the Lagrange multipliers to express the multipliers in terms of the original variables.

As early as 1977, Mang and Gallagher (1977) reported on the critical assessment of a typical SVP, in the context of the finite element method, denoted as "simplified hybrid displacement method." They have shown that this simplified variational principle is inadequate. They examined its use for satisfaction of the relaxed C^1-interelement continuity requirement for plate-bending finite elements based on Kirchhoff's theory. It was shown that this method does not serve its purpose, which is enforcement of the relaxed C^1-interelement continuity, at least in a weak sense, with the help of algebraic constraint equations. In 1985, Mang, Hofstetter, and Gallagher (1985) demonstrated several deficiencies of SVPs. This reference contains a comprehensive literature review.

In 1985, Xue and Atluri (1985) examined mixed-hybrid finite element methods based on multi-field variational principles with regards to the conditions for the existence, stability, and uniqueness of their solutions. They discussed the reasons as to why certain "simplified hybrid-mixed methods," in general, and the "simplified hybrid displacement method," in particular, become unstable. Their mathematical proofs are restricted to linear solid mechanics. No numerical results were presented.

The impact of the criticism raised in Mang and Gallagher (1977), Mang, Hofstetter, and Gallagher (1985), and Xue and Atluri (1985) was relatively small. The main reason for this seems to have been the lack of an analytical proof of the existence of fundamental shortcomings of SVPs without regards to finite element discretizations and without the restriction to linearity. Apparently, the deficiencies of these variational principles were attributed primarily to inadequate, i.e., too coarse, finite element discretizations.

The purpose of this paper is to prove that the aforementioned conversion of a LMM to the corresponding SVP may lead to the loss of uniqueness of solutions of boundary value problems. (The term "uniqueness of a solution" is defined as the absence of solutions in the variational neighbourhood of this solution.) A particularly simple form of the geometrically nonlinear theory of bending of beams is used as the vehicle for this proof.

The theoretical part of the paper starts with the formulation of the LMM for a beam subjected to geometric boundary conditions only, all of which have been relaxed. With the help of the resulting Euler-Lagrange equations the corresponding SVP is formulated. On the basis of the Trefftz condition it is shown that special variations of the displacements exist (where they should not exist), for which the second variation of the underlying energy functional is stationary. Hence, in contrast to the LMM-based solution, the SVP-based solution of the investigated problem is not unique.

A classical buckling problem is chosen as an example for demonstrating the deficiencies of the SVP. It is shown that application of this variational principle results in eigenvalues in addition to the correct ones. In particular, the smallest eigenvalue, which is related to the buckling load, is obtained as one half of the correct value.

2 Theoretical Investigation

2.1 Lagrange Multiplier Method. Figure 1 shows a beam with a variable cross section subjected to a nonuniformly distributed vertical load $\bar{p}$. All boundary conditions are geometric boundary conditions. They are given as

$$u_0 = \overline{u_0}, \quad w_0 = \overline{w_0}, \quad \varphi_0 = \overline{\varphi_0}, \tag{1.1}$$

$$u_l = \overline{u_l}, \quad w_l = \overline{w_l}, \quad \varphi_l = \overline{\varphi_l}, \tag{1.2}$$

where u and w denote the horizontal and the vertical displacement, respectively, and $\varphi = dw/dx$ stands for the slope of the deflection curve. $u_0 = u(x = 0)$, $u_l = u(x = l)$, etc., with l

[1] This paper is dedicated to Prof. E. Stein on the occasion of his 65th birthday.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Apr. 11, 1995; final revision, Nov. 22, 1995. Associate Technical Editor: J. N. Reddy.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

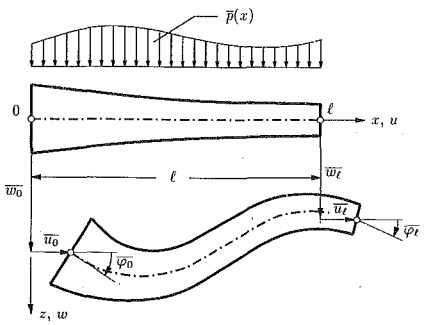


Fig. 1 Beam with a variable cross section subjected to a nonuniformly distributed vertical load; inhomogeneous geometric boundary conditions

representing the length of the beam. The overbars refer to prespecified quantities.

Assuming that all boundary conditions have been relaxed, the functional that serves as the starting point for determination of the deflection curve takes on the following form:

$$\Pi_\lambda = \Pi - \lambda_1(\varphi_0 - \overline{\varphi_0}) + \lambda_2(\varphi_l - \overline{\varphi_l}) + \lambda_3(w_0 - \overline{w_0}) - \lambda_4(w_l - \overline{w_l}) + \lambda_5(u_0 - \overline{u_0}) - \lambda_6(u_l - \overline{u_l}) \rightarrow \text{stationary.} \quad (2)$$

In Eq. (2), Π is the potential energy, given as

$$\Pi = U + V, \quad (3)$$

where U is the strain energy, obtained as

$$U = \tfrac{1}{2}\int_0^l (\mathrm{EI}\kappa^2 + \mathrm{EA}\epsilon^2)dx, \quad (4)$$

with κ, ϵ, EI, and EA denoting, in this order, the curvature of the deflection curve, the extension of the axis of the beam, the flexural rigidity, and the extensional rigidity. On the basis of the geometrically nonlinear theory of small displacements and moderately large rotations, κ and ϵ are given as

$$\kappa = -\frac{d^2w}{dx^2}, \quad \epsilon = \frac{du}{dx} + \frac{1}{2}\varphi^2. \quad (5)$$

V is the load potential, defined as

$$V = -\int_0^l \bar{p}w dx. \quad (6)$$

The quantities $\lambda_1, \lambda_2, \ldots \lambda_6$ in Eq. (2) are the Lagrange multipliers, introduced to enforce the geometric boundary conditions.

The condition for Π_λ to take on a stationary value is

$$\delta\Pi_\lambda = 0, \quad (7)$$

where $\delta\Pi_\lambda$ denotes the variation of Π_λ. $\delta\Pi_\lambda$ is obtained as

$$\delta\Pi_\lambda = \int_0^l (\mathrm{EI}\kappa\delta\kappa + \mathrm{EA}\epsilon\delta\epsilon - \bar{p}\delta w)dx - \delta\lambda_1(\varphi_0 - \overline{\varphi_0}) - \lambda_1\delta\varphi_0 + \delta\lambda_2(\varphi_l - \overline{\varphi_l}) + \lambda_2\delta\varphi_l + \delta\lambda_3(w_0 - \overline{w_0}) + \lambda_3\delta w_0 - \delta\lambda_4(w_l - \overline{w_l}) - \lambda_4\delta w_l + \lambda_5\delta u_0 + \delta\lambda_5(u_0 - \overline{u_0}) - \lambda_6\delta u_l - \delta\lambda_6(u_l - \overline{u_l}). \quad (8)$$

Integration by parts of the first and the second term of the integral in Eq. (8) yields

$$\int_0^l \mathrm{EI}\kappa\delta\kappa dx = -\mathrm{EI}\kappa\delta\varphi\big|_0^l + \int_0^l \frac{d}{dx}(\mathrm{EI}\kappa)\delta\varphi dx = -\mathrm{EI}\kappa\delta\varphi\big|_0^l + \frac{d}{dx}(\mathrm{EI}\kappa)\delta w\big|_0^l - \int_0^l \frac{d^2}{dx^2}(\mathrm{EI}\kappa)\delta w dx, \quad (9)$$

$$\int_0^l \mathrm{EA}\epsilon\delta\epsilon dx = \int_0^l \mathrm{EA}\epsilon\delta\left(\frac{du}{dx} + \frac{1}{2}\varphi^2\right)dx = \int_0^l \mathrm{EA}\epsilon\left[\delta\left(\frac{du}{dx}\right) + \varphi\delta\varphi\right]dx = \mathrm{EA}\epsilon(\delta u + \varphi\delta w)\Big|_0^l - \int_0^l \left[\frac{d}{dx}(\mathrm{EA}\epsilon)\delta u + \frac{d}{dx}(\mathrm{EA}\epsilon\varphi)\delta w\right]dx. \quad (10)$$

Substitution of the Eqs. (9) and (10) into Eq. (8), and rearrangement, results in

$$\delta\Pi_\lambda = -\int_0^l \left\{\left[\frac{d^2}{dx^2}(\mathrm{EI}\kappa) + \frac{d}{dx}(\mathrm{EA}\epsilon\varphi) + \bar{p}\right]\delta w + \frac{d}{dx}(\mathrm{EA}\epsilon)\delta u\right\}dx - \delta\lambda_1(\varphi_0 - \overline{\varphi_0}) + [(\mathrm{EI}\kappa)_0 - \lambda_1]\delta\varphi_0 + \delta\lambda_2(\varphi_l - \overline{\varphi_l}) - [(\mathrm{EI}\kappa)_l - \lambda_2]\delta\varphi_l + \delta\lambda_3(w_0 - \overline{w_0}) - \left\{\left[\frac{d}{dx}(\mathrm{EI}\kappa)\right]_0 + (\mathrm{EA}\epsilon\varphi)_0 - \lambda_3\right\}\delta w_0 - \delta\lambda_4(w_l - \overline{w_l}) + \left\{\left[\frac{d}{dx}(\mathrm{EI}\kappa)\right]_l + (\mathrm{EA}\epsilon\varphi)_l - \lambda_4\right\}. \quad (11)$$

Because of the independence of the variations in Eq. (11), the condition $\delta\Pi_\lambda = 0$ yields

$$-\frac{d^2}{dx^2}(\mathrm{EI}\kappa) - \frac{d}{dx}(\mathrm{EA}\epsilon\varphi) = \bar{p}, \quad (12.1)$$

$$\frac{d}{dx}(\mathrm{EA}\epsilon) = 0, \quad (12.2)$$

$$\varphi_0 = \overline{\varphi_0}, \quad \varphi_l = \overline{\varphi_l}, \quad w_0 = \overline{w_0}, \quad w_l = \overline{w_l}, \quad u_0 = \overline{u_0}, \; u_l = \overline{u_l}, \quad (12.3)$$

$$\lambda_1 = (\mathrm{EI}\kappa)_0, \quad \lambda_2 = (\mathrm{EI}\kappa)_l, \quad \lambda_3 = \left[\frac{d}{dx}(\mathrm{EI}\kappa)\right]_0 + (\mathrm{EA}\epsilon\varphi)_0, \quad \lambda_4 = \left[\frac{d}{dx}(\mathrm{EI}\kappa)\right]_l + (\mathrm{EA}\epsilon\varphi)_l,$$

$$\lambda_5 = (\mathrm{EA}\epsilon)_0, \quad \lambda_6 = (\mathrm{EA}\epsilon)_l. \quad (12.4)$$

The Eqs. (12) are the Euler-Lagrange equations of the variational principle $\Pi_\lambda \rightarrow$ *stationary.* The Eqs. (12.1) and (12.2) are the field equations, the Eqs. (12.3) are the geometric boundary conditions, and the relations (12.4) are the equations of definition for the Lagrange multipliers. They represent the bending moments, the transverse forces, and the axial forces at the two ends of the beam.

2.2 Simplified Variational Principle. Inserting the Eqs. (12.4) into the expression for Π_λ (Eq. (2)) results in the following functional (Π_S) of the Simplified Variational Principle:

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

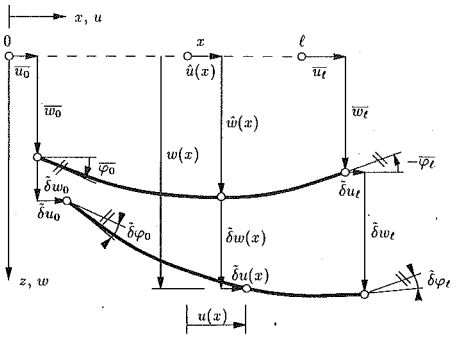


Fig. 2 True solution and solution in its variational neighborhood satisfying the field equations but violating the relaxed geometric boundary conditions

$$\Pi_S = \Pi + \left\{ \mathrm{EI}\kappa(\varphi - \bar{\varphi}) - \left[\frac{d}{dx}(\mathrm{EI}\kappa) + \mathrm{EA}\epsilon\varphi \right](w - \bar{w}) - \mathrm{EA}\epsilon(u - \bar{u}) \right\}\Big|_0^l \rightarrow \text{stationary.} \quad (13)$$

The condition for Π_S to take on a stationary value is $\delta\Pi_S = 0$. This gives

$$\delta\Pi_S = -\int_0^l \left\{ \left[\frac{d^2}{dx^2}(\mathrm{EI}\kappa) + \frac{d}{dx}(\mathrm{EA}\epsilon\varphi) + \bar{p} \right]\delta w + \frac{d}{dx}(\mathrm{EA}\epsilon)\delta u \right\} dx + \left\{ \mathrm{EI}\delta\kappa(\varphi - \bar{\varphi}) - \left[\frac{d}{dx}(\mathrm{EI}\delta\kappa) + \mathrm{EA}\delta(\epsilon\varphi) \right](w - \bar{w}) - \mathrm{EA}\delta\epsilon(u - \bar{u}) \right\}\Big|_0^l = 0. \quad (14)$$

Because of the independence of the variations δu and δw, this yields the field Eqs. (12.1) and (12.2). Noting, from (12.2), that $\mathrm{EA}\epsilon$ is constant, say C, and $(d\varphi/dx) = -\kappa$, Eq. (12.1) is

$$-\frac{d^2}{dx^2}(\mathrm{EI}\kappa) + C\kappa = \bar{p}. \quad (15)$$

2.3 Nonuniqueness of the SVP-Based Solution of the Given Problem. It will now be shown that the Simplified Variational Principle does not allow satisfaction of the relaxed geometric boundary conditions. To do so, it will be demonstrated that, in any case, the SVP-based solution of the given problem will be found to be nonunique. This nonuniqueness is characterized by

$$u = \hat{u} + \tilde{\delta}u, \quad w = \hat{w} + \tilde{\delta}w, \quad (16)$$

where $\hat{u}$, $\hat{w}$ symbolizes the true solution and $\tilde{\delta}u$, $\tilde{\delta}w$ denotes special variations such that the stationary condition

$$\delta\Pi_S(u, w) = \delta\Pi_S(\hat{u} + \tilde{\delta}u, \hat{w} + \tilde{\delta}w) = 0 \quad (17)$$

is satisfied for nonvanishing $\tilde{\delta}u$ and/or $\tilde{\delta}w$. (With the exception of bifurcation points on load-displacement paths, such variations must not exist.)

Figure 2 shows the true solution and a solution in its variational neighborhood. The latter does not satisfy the geometric boundary conditions. Hence,

$$u_0 - \overline{u_0} = \tilde{\delta}u_0, \quad w_0 - \overline{w_0} = \tilde{\delta}w_0, \quad \varphi_0 - \overline{\varphi_0} = \tilde{\delta}\varphi_0,$$
$$u_l - \overline{u_l} = \tilde{\delta}u_l, \quad w_l - \overline{w_l} = \tilde{\delta}w_l, \quad \varphi_l - \overline{\varphi_l} = \tilde{\delta}\varphi_l. \quad (18)$$

It follows from the Eqs. (14), (15), (17), and (18) that

$$-\frac{d^2}{dx^2}[\mathrm{EI}(\hat{\kappa} + \tilde{\delta}\kappa)] + (\hat{C} + \tilde{\delta}C)(\hat{\kappa} + \tilde{\delta}\kappa) = \bar{p}, \quad (19)$$

$$\mathrm{EA}(\hat{\epsilon} + \tilde{\delta}\epsilon) = \hat{C} + \tilde{\delta}C, \quad (20)$$

$$\left\langle \mathrm{EI}\delta(\hat{\kappa} + \tilde{\delta}\kappa)\tilde{\delta}\varphi - \left\{ \frac{d}{dx}[\mathrm{EI}\delta(\hat{\kappa} + \tilde{\delta}\kappa)] + \mathrm{EA}\delta[(\hat{\epsilon} + \tilde{\delta}\epsilon)(\hat{\varphi} + \tilde{\delta}\varphi)] \right\}\tilde{\delta}w - \mathrm{EA}\delta(\hat{\epsilon} + \tilde{\delta}\epsilon)\tilde{\delta}u \right\rangle\Big|_0^l = 0, \quad (21)$$

where the relations (19) and (20) are the field equations for solutions in the variational neighbourhood of the true solution, whereas Eq. (21) refers to the boundary values of these solutions.

Subtracting Eq. (15) from Eq. (19) and neglecting the term $\tilde{\delta}C\tilde{\delta}\kappa$ because it is of higher order small yields

$$\frac{d^2}{dx^2}(\mathrm{EI}\tilde{\delta}\kappa) - \hat{C}\tilde{\delta}\kappa - \hat{\kappa}\tilde{\delta}C = 0. \quad (22)$$

Also, since $\mathrm{EA}\hat{\epsilon} = \hat{C}$, Eq. (20) results in

$$\mathrm{EA}\tilde{\delta}\epsilon = \tilde{\delta}C. \quad (23)$$

According to Fig. 3,

$$\delta u = \delta(\hat{u} + \tilde{\delta}u) = \delta(\tilde{\delta}u), \quad \delta w = \delta(\hat{w} + \tilde{\delta}w) = \delta(\tilde{\delta}w). \quad (24)$$

Based on the relations (24) and on neglecting the term $\{\mathrm{EA}[\delta(\tilde{\delta}\epsilon)\tilde{\delta}\varphi + \delta(\tilde{\delta}\varphi)\tilde{\delta}\epsilon]\tilde{\delta}w\}\,|_0^l$, which is small of higher order, Eq. (21) can be rewritten as

$$\left\langle \mathrm{EI}\delta(\tilde{\delta}\kappa)\tilde{\delta}\varphi - \left\{ \frac{d}{dx}[\mathrm{EI}\delta(\tilde{\delta}\kappa)] + \mathrm{EA}[\hat{\epsilon}\delta(\tilde{\delta}\varphi) + \hat{\varphi}\delta(\tilde{\delta}\epsilon)] \right\}\tilde{\delta}w - \mathrm{EA}\delta(\tilde{\delta}\epsilon)\tilde{\delta}u \right\rangle\Big|_0^l = 0. \quad (25)$$

The Eqs. (22), (23), and (25) could have been derived directly from the Trefftz condition

$$\delta(\tilde{\delta}^2\Pi_S) = 0, \quad (26)$$

expressing the fact that $\delta^2\Pi_S$ is stationary for the special variations $\tilde{\delta}u$, $\tilde{\delta}w$.

Making use of $\mathrm{EA}\hat{\epsilon} = \hat{C}$ and of Eq. (23), Eq. (26) can be rewritten as

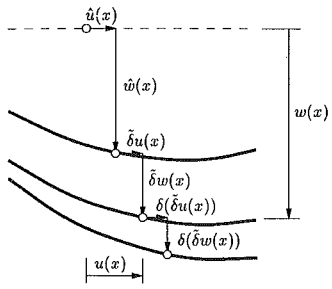


Fig. 3 General variations $\delta(\tilde{\delta}u(x))$, $\delta(\tilde{\delta}w(x))$ of special variations $\tilde{\delta}u(x)$, $\tilde{\delta}w(x)$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\left\langle \mathrm{EI}\delta(\tilde{\delta}\kappa)\tilde{\delta}\varphi - \left\{\frac{d}{dx}\left[\mathrm{EI}\delta(\tilde{\delta}\kappa)\right] + \hat{C}\delta(\tilde{\delta}\varphi) + \hat{\varphi}\delta(\tilde{\delta}C)\right\}\tilde{\delta}w - \delta(\tilde{\delta}C)\tilde{\delta}u\right\rangle\Big|_0^l = 0. \quad (27)$$

Integration of Eq. (22) yields

$$\frac{d}{dx}(\mathrm{EI}\tilde{\delta}\kappa) + \hat{C}\tilde{\delta}\varphi + \hat{\varphi}\tilde{\delta}C + \tilde{\delta}D = 0, \quad (28)$$

where $\tilde{\delta}D$ is an integration constant. Its variation is

$$\frac{d}{dx}\left[\mathrm{EI}\delta(\tilde{\delta}\kappa)\right] + \hat{C}\delta(\tilde{\delta}\varphi) + \hat{\varphi}\delta(\tilde{\delta}C) + \delta(\tilde{\delta}D) = 0. \quad (29)$$

Substitution of Eq. (29) into (27) gives

$$\left[\mathrm{EI}\delta(\tilde{\delta}\kappa)\tilde{\delta}\varphi + \delta(\tilde{\delta}D)\tilde{\delta}w - \delta(\tilde{\delta}C)\tilde{\delta}u\right]\Big|_0^l = 0. \quad (30)$$

In order to obtain expressions for the boundary values of $\delta(\tilde{\delta}\kappa)$ it is necessary to distinguish between the general case, $\hat{C} \neq 0$, which refers to a nonvanishing axial force, and the special case $\hat{C} = 0$, which refers to zero axial force.

2.4 $\hat{C} \neq 0$ (Nonvanishing Axial Force): Equation (22) can be rewritten as

$$\frac{d^2}{dx^2}(\tilde{\delta}\kappa) + \frac{\frac{d^2}{dx^2}(\mathrm{EI}) - \hat{C}}{\mathrm{EI}}\tilde{\delta}\kappa = \frac{\hat{\kappa}}{\mathrm{EI}}\tilde{\delta}C. \quad (31)$$

The general solution of this nonhomogeneous linear differential equation of second order can formally be written as

$$\tilde{\delta}\kappa = \tilde{\delta}\kappa_h + \tilde{\delta}\kappa_p, \quad (32)$$

where $\tilde{\delta}\kappa_h$ is a general solution of the respective homogeneous differential equations and $\tilde{\delta}\kappa_p$ is a particular solution of the nonhomogeneous differential equation. $\tilde{\delta}\kappa_h$ is obtained as

$$\tilde{\delta}\kappa_h = y_1\tilde{\delta}A + y_2\tilde{\delta}B, \quad (33)$$

where $y_1 = y_1(x)$, $y_2 = y_2(x)$ and $\tilde{\delta}A$ and $\tilde{\delta}B$ are constants. With the help of the method of variation of parameters $\tilde{\delta}\kappa_p$ is obtained as

$$\tilde{\delta}\kappa_p = \left[-y_1(x)\int\frac{y_2(\xi)r(\xi)}{W(\xi)}d\xi + y_2(x)\int\frac{y_1(\xi)r(\xi)}{W(\xi)}d\xi\right]\tilde{\delta}C, \quad (34)$$

where $r = r(\xi)$ is $r = \hat{\kappa}/\mathrm{EI}$, and $W = W(\xi)$ is the Wronskian, given as

$$W = y_1\frac{dy_2}{d\xi} - y_2\frac{dy_1}{d\xi}. \quad (35)$$

Thus,

$$\tilde{\delta}\kappa = y_1\tilde{\delta}A + y_2\tilde{\delta}B + y_p\tilde{\delta}C, \quad (36)$$

where $y_p = y_p(x)$ is an abbreviation, obtained as

$$y_p = -y_1\int\frac{y_2 r}{W}d\xi + y_2\int\frac{y_1 r}{W}d\xi. \quad (37)$$

Hence,

$$\delta(\tilde{\delta}\kappa) = y_1\delta(\tilde{\delta}A) + y_2\delta(\tilde{\delta}B) + y_p\delta(\tilde{\delta}C). \quad (38)$$

Substitution of Eq. (38) into Eq. (30) yields

$$\left\{\mathrm{EI}[y_1\delta(\tilde{\delta}A) + y_2\delta(\tilde{\delta}B) + y_p\delta(\tilde{\delta}C)]\tilde{\delta}\varphi + \delta(\tilde{\delta}D)\tilde{\delta}w - \delta(\tilde{\delta}C)\tilde{\delta}u\right\}\Big|_0^l = 0. \quad (39)$$

$\delta(\tilde{\delta}A)$, $\delta(\tilde{\delta}B)$, $\delta(\tilde{\delta}C)$, and $\delta(\tilde{\delta}D)$ are arbitrary variations. Consequently,

$$\boxed{(\mathrm{EI}y_1\tilde{\delta}\varphi)\big|_0^l = 0, \quad (\mathrm{EI}y_2\tilde{\delta}\varphi)\big|_0^l = 0, \quad (\mathrm{EI}y_p\tilde{\delta}\varphi - \tilde{\delta}u)\big|_0^l = 0, \quad \tilde{\delta}w\big|_0^l = 0.} \quad (40)$$

The Eqs. (40) represent a system of four homogeneous algebraic equations for the six special variations $\tilde{\delta}u_0$, $\tilde{\delta}u_l$, $\tilde{\delta}w_0$, $\tilde{\delta}w_l$, $\tilde{\delta}\varphi_0$, $\tilde{\delta}\varphi_l$. *Hence, the SVP-based solution of the given problem is not unique. Q.E.D.*

In Appendix A this circumstance is developed in detail. It begins with integration of Eq. (36) and proceeds to the algebraic representation of Eq. (40). The solution of that representation demonstrates that incorrect eigenforms and eigenvalues are obtained for the subject problem in addition to the correct ones.

2.5 $\hat{C} = 0$ (Vanishing Axial Force): Specialization of Eq. (29) for $\hat{C} = 0$ and integration of the resulting relation yields

$$\mathrm{EI}\delta(\tilde{\delta}\kappa) + \hat{w}\delta(\tilde{\delta}C) + x\delta(\tilde{\delta}D) + \delta(\tilde{\delta}E) = 0, \quad (41)$$

where $\delta(\tilde{\delta}E)$ is an integration constant. Substitution of Eq. (41) into Eq. (30) gives

$$\left\{-[\hat{w}\delta(\tilde{\delta}C) + x\delta(\tilde{\delta}D) + \delta(\tilde{\delta}E)]\tilde{\delta}\varphi + \delta(\tilde{\delta}D)\tilde{\delta}w - \delta(\tilde{\delta}C)\tilde{\delta}u\right\}\Big|_0^l = 0. \quad (42)$$

$\delta(\tilde{\delta}C)$, $\delta(\tilde{\delta}D)$, and $\delta(\tilde{\delta}E)$ are arbitrary variations. Consequently,

$$\boxed{(\hat{w}\tilde{\delta}\varphi + \tilde{\delta}u)\big|_0^l = 0, \quad (\tilde{\delta}w - x\tilde{\delta}\varphi)\big|_0^l = 0, \quad \tilde{\delta}\varphi\big|_0^l = 0.} \quad (43)$$

The Eqs. (43) represent a system of three homogeneous linear algebraic equations for the six special variations $\tilde{\delta}u_0$, $\tilde{\delta}u_l$, $\tilde{\delta}w_0$, $\tilde{\delta}w_l$, $\tilde{\delta}\varphi_0$, $\tilde{\delta}\varphi_l$. Hence, expectedly, also for the special case $\hat{C} = 0$ the SVP-based solution of the given problem is not unique.

Appendix B demonstrates the algebraic formation of Eqs. (43) and their solution. Again, it is shown that incorrect eigenforms and eigenvalues exist in addition to the correct ones.

2.6 Static Boundary Condition. In this Section the geometric boundary condition $u_l = \overline{u_l}$ shown in Fig. 1 will be replaced by the static boundary condition $N = \bar{P}$, where $\bar{P}$ is an axial force acting at a sliding support at $x = l$. Hence, the term $-\lambda_6(u_l - \overline{u_l})$ in Eq. (2) must be replaced by the term $-\bar{P}u_l$. Accordingly,

$$\mathrm{EA}\hat{\epsilon} = \hat{C} = \bar{P} \quad (44)$$

and

$$\mathrm{EA}\tilde{\delta}\epsilon = \tilde{\delta}C = 0. \quad (45)$$

Thus,

$$\delta(\tilde{\delta}C) = 0. \quad (46)$$

If $\bar{P} = \hat{C} \neq 0$, then, following from Eq. (39),

$$(\mathrm{EI}y_1\tilde{\delta}\varphi)\big|_0^l = 0, \quad (\mathrm{EI}y_2\tilde{\delta}\varphi)\big|_0^l = 0, \quad \tilde{\delta}w\big|_0^l = 0. \quad (47)$$

Hence, Eq. (40.3) no longer exists. Nevertheless, according to Eq. (A.5), $\tilde{\delta}u$ is, at best, only determined up to a constant.

Analogous to expressing the Eqs. (40) in terms of $\tilde{\delta}A$, $\tilde{\delta}B$, $\tilde{\delta}C$, and $\tilde{\delta}D$, the Eqs. (47) can be expressed in terms of $\tilde{\delta}A$, $\tilde{\delta}B$, and $\tilde{\delta}D$. For this purpose the relations (A.1) and (A.2) must be specialized for $\tilde{\delta}C = 0$.

If $\bar{P} = \hat{C} = 0$, then, following from Eq. (42),

$$(\tilde{\delta}w - x\tilde{\delta}\varphi)\big|_0^l = 0, \ \tilde{\delta}\varphi\big|_0^l = 0. \quad (48)$$

(These equations are identical to the relations (12.1) and (12.2)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

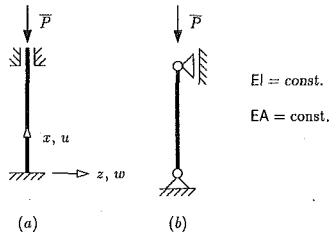


Fig. 4 Beam, clamped at both ends, subjected to an axial point load at the upper end: (*a*) actual boundary conditions, (*b*) geometric boundary conditions only partially relaxed

in Xue and Atluri (1985). Following from Eq. (12.3) in Xue and Atluri (1985), the Eqs. (12.1) and (12.2) require that $\tilde{C} = 0$, where $\tilde{C}$ is the part of $\hat{C}$ resulting from the nonlinear part of $\hat{\epsilon}$. In this context it should be recalled that the linear part of $\hat{\epsilon}$ was a priori neglected in Xue and Atluri (1985). The vanishing of the horizontal force $\hat{C}$ is in conflict with the tacit assumption of $\overline{u_0} = \overline{u_l} = 0$ in Xue and Atluri (1985)).

Hence, Eq. (43.1) no longer exists. Nevertheless, according to Equation (A.5), $\tilde{\delta}u$ is, at best, only determined up to a constant.

Analogous to expressing the Eqs. (43) in terms of $\tilde{\delta}C$, $\tilde{\delta}D$, $\tilde{\delta}E$, and $\tilde{\delta}G$, the Eqs. (48) can be expressed in terms of $\tilde{\delta}D$, $\tilde{\delta}E$, and $\tilde{\delta}G$. For this purpose the relations (B.2) and (B.3) must be specialized for $\tilde{\delta}C = 0$.

In any case, the SVP-based solution is not unique.

3 Example

In order to demonstrate the deficiencies of the SVP, the buckling load of a beam with constant flexural and extensional rigidity, clamped at both ends and subjected to an axial point load at the upper end (Fig. 4(*a*)), is determined on the basis of this variational principle.

Substitution of $N = \hat{C} = -\bar{P}$ and of Eq. (45) into Eq. (22) and consideration of EI = const. yields

$$\frac{d^4}{dx^4}(\tilde{\delta}w) + \alpha^2 \frac{d^2}{dx^2}(\tilde{\delta}w) = 0, \tag{49}$$

where $\alpha^2 = \bar{P}/\mathrm{EI}$ and $\overline{\delta\kappa} = -(d^2/dx^2)(\overline{\delta w})$. (According to Eq. (31) the same differential equation would be obtained for $\bar{P} = 0$ and $\mathrm{EI} = (\mathrm{EI})_0 e^{\alpha x}$, where $(\mathrm{EI})_0 = \mathrm{EI}|_{x=0}$. Hence, even for $\bar{P} = 0$, the variations $\tilde{\delta}u$ and $\tilde{\delta}w$ are not necessarily restricted to a rigid-body motion.)

The general solution of this homogeneous linear differential equation of fourth order is obtained as

$$\tilde{\delta}w = \cos\alpha x\, \tilde{\delta}A + \sin\alpha x\, \tilde{\delta}B + x\tilde{\delta}D + \tilde{\delta}E. \tag{50}$$

Thus,

$$\tilde{\delta}\varphi = -\alpha \sin\alpha x\, \tilde{\delta}A + \alpha \cos\alpha x\, \tilde{\delta}B + \tilde{\delta}D, \tag{51}$$

$$\tilde{\delta}\kappa = \alpha^2 \cos\alpha x\, \tilde{\delta}A + \alpha^2 \sin\alpha x\, \tilde{\delta}B. \tag{52}$$

Comparing Eq. (52) with Eq. (36) shows that

$$y_1 = \alpha^2 \cos\alpha x, \quad y_2 = \alpha^2 \sin\alpha x. \tag{53}$$

The boundary values of $\tilde{\delta}w$, $\tilde{\delta}\varphi$, y_1, and y_2 are obtained as follows:

$$\tilde{\delta}w_0 = \tilde{\delta}A + \tilde{\delta}E,$$

$$\tilde{\delta}w_l = \cos\alpha l\, \tilde{\delta}A + \sin\alpha l\, \tilde{\delta}B + l\tilde{\delta}D + \tilde{\delta}E, \tag{54}$$

$$\tilde{\delta}\varphi_0 = \alpha\tilde{\delta}B + \tilde{\delta}D, \tag{55}$$

$$\tilde{\delta}\varphi_l = -\alpha \sin\alpha l\, \tilde{\delta}A + \alpha \cos\alpha l\, \tilde{\delta}B + \tilde{\delta}D,$$

$$y_1|_0 = \alpha^2, \quad y_1|_l = \alpha^2 \cos\alpha l, \tag{56}$$

$$y_2|_0 = 0, \quad y_2|_l = \alpha^2 \sin\alpha l. \tag{57}$$

Making use of these relations in the Eqs. (47) and considering EI = const. gives

$$(y_1\tilde{\delta}\varphi)|_0^l = \alpha^2 \cos\alpha l(-\alpha \sin\alpha l\, \tilde{\delta}A + \alpha \cos\alpha l\, \tilde{\delta}B + \tilde{\delta}D) - \alpha^2(\alpha\tilde{\delta}B + \tilde{\delta}D) = 0, \tag{58}$$

$$(y_2\tilde{\delta}\varphi)|_0^l = \alpha^2 \sin\alpha l(-\alpha \sin\alpha l\, \tilde{\delta}A + \alpha \cos\alpha l\, \tilde{\delta}B + \tilde{\delta}D) = 0, \tag{59}$$

$$\tilde{\delta}w|_0^l = (\cos\alpha l - 1)\tilde{\delta}A + \sin\alpha l\, \tilde{\delta}B + l\tilde{\delta}D = 0. \tag{60}$$

If the geometric boundary condition $w_0 - \overline{w_0} = 0$ was not relaxed, i.e., if $\tilde{\delta}w_0 = 0$, then, following from Eq. (54.1),

$$\tilde{\delta}E = -\tilde{\delta}A. \tag{61}$$

Substitution of Eq. (61) into Eq. (54.2) gives

$$\tilde{\delta}w_l = (\cos\alpha l - 1)\tilde{\delta}A + \sin\alpha l\, \tilde{\delta}B + l\tilde{\delta}D. \tag{62}$$

According to Eq. (60), the above expression for $\tilde{\delta}w_l$ is equal to zero. Hence, in the given case, satisfaction of Eq. (60) implies satisfaction of the boundary conditions $\tilde{\delta}w_0 = 0$ and $\tilde{\delta}w_l = 0$. Consequently, the Eqs. (58), (59), and (60) would also be obtained for a beam, clamped at both ends, for which only the geometric boundary conditions $\tilde{\delta}\varphi_0 = 0$ and $\tilde{\delta}\varphi_l = 0$ were relaxed (Fig. 4(*b*).

After division of the Eqs. (58) and (59) by α^2, the Eqs. (58), (59), and (60) are rewritten as

$$\sin\alpha l \begin{bmatrix} \alpha \sin\alpha l \cos\alpha l & \alpha(1 - \cos^2\alpha l) & (1 - \cos\alpha l) \\ -\alpha \sin\alpha l & \alpha \cos\alpha l & 1 \\ (1 - \cos\alpha l) & -\sin\alpha l & -l \end{bmatrix} \times \begin{Bmatrix} \tilde{\delta}A \\ \tilde{\delta}B \\ \tilde{\delta}D \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}. \tag{63}$$

This system of homogeneous linear algebraic equations will have a nontrivial solution if and only if

$$\sin\alpha l\, \mathrm{Det} \begin{bmatrix} \alpha \sin\alpha l \cos\alpha l & \alpha(1 - \cos^2\alpha l) & (1 - \cos\alpha l) \\ -\alpha \sin\alpha l & \alpha \cos\alpha l & 1 \\ (1 - \cos\alpha l) & -\sin\alpha l & -l \end{bmatrix} = 4\alpha \sin\alpha l \left(\sin\frac{\alpha l}{2} - \frac{\alpha l}{2}\cos\frac{\alpha l}{2}\right)\sin\frac{\alpha l}{2} = 0. \tag{64}$$

Equation (64) will be satisfied if

(*a*) $\sin\alpha l = 0 \rightarrow \alpha l = m\pi, \quad m = 1, 2, \ldots \infty,$ or (65.1)

(*b*) $\sin\frac{\alpha l}{2} = 0 \rightarrow \alpha l = 2n\pi, \quad n = 1, 2, \ldots \infty,$ or (65.2)

(*c*) $\sin\frac{\alpha l}{2} - \frac{\alpha l}{2}\cos\frac{\alpha l}{2} = 0 \rightarrow \alpha l = (2n + 1 - \beta_n)\pi,$

$n = 1, 2, \ldots \infty, \beta_1 = 0, 139 \ldots, \ldots \underset{n\to\infty}{\beta_n} \rightarrow 0.$ (65.3)

Equation (65.1) is the characteristic equation of a simply supported beam, subjected to an axial point load at one end. Hence, the respective eigenvalues represent incorrect solutions of the actual problem. In contrast to these eigenvalues, the ones obtained from the equations (65.2) and (65.3) are the correct eigenvalues for a clamped beam, subjected to an axial point load at one end. For $m = 2n$, $n = 1, 2, \ldots \infty$, the Eqs. (65.1) and (65.2) are satisfied simultaneously. Consequently, the respective eigenvalues are double eigenvalues.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

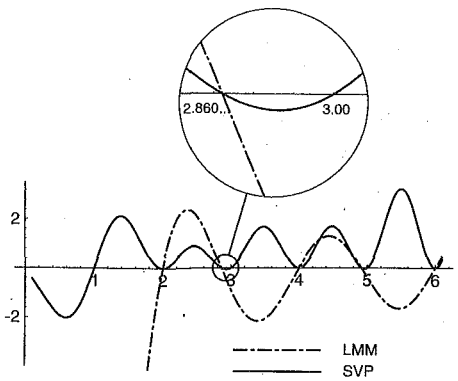


Fig. 5 Scaled function according to Eq. (64) (SVP) and to the corresponding true solution (LMM). The values along the abscissa are $\alpha l/\pi$.

Figure 5 illustrates the suitably scaled expression on the left-hand side of the SVP-based Eq. (64). The figure also shows the corresponding expression resulting from the Lagrange Multiplier Method. It is seen that application of the SVP leads to eigenvalues in addition to the correct ones. In particular, the smallest eigenvalue is obtained as one half of the correct value. Hence, the value for the buckling load is obtained as one quarter of the correct value.

4 Conclusions

It was shown that the conversion of a Lagrange Multiplier Method to the corresponding Simplified Variational Principle has led to the loss of uniqueness of solutions of boundary value problems. For this purpose a beam with a variable cross section subjected to a nonuniformly distributed vertical load was investigated. Initially, only geometric boundary conditions, all of which were relaxed, were considered. By means of the geometrically nonlinear theory of small displacements and moderately large rotations it was demonstrated that the SVP, in contrast to the LMM, was not capable of enforcing the relaxed boundary conditions or, at least, not all of them. The SVP-based solution is only determined up to rigid-body displacements of the beam in each one of the two coordinate directions.

If, incidentally, the constant axial force is zero, the SVP-based solution will only be determined up to a rigid-body motion consisting of rigid-body displacements in the two coordinate directions and of a rigid-body rotation. It was also shown that the replacement of the geometric boundary condition for the displacement component in the axial direction at one of the two ends of the beam by the respective inhomogeneous or homogeneous static boundary condition has no essential influence on the described situation.

In the context of Euler buckling of a clamped beam it was shown that incorrect eigenvalues exist in addition to the correct ones. In particular, the smallest eigenvalue, which is related to the buckling load, was obtained as one half of the correct value.

The investigation reported in this paper was entirely analytical. Hence, the deficiencies of SVP-based solutions of boundary value problems are fundamental. Consequently, the shortcomings of numerical solutions on the basis of the finite element method, which were demonstrated in previous papers, cannot be attributed primarily to inadequate, i.e., too coarse finite element discretizations.

References

Mang, H. A., and Gallagher, R. H., 1977, "A Critical Assessment of the Simplified Hybrid Displacement Method," *Int. J. Num. Meth. Engng.* Vol. 11, pp. 145–167.

Mang, H. A., and Helnwein, P., 1993, "Zur Untauglichkeit vereinfachter Variationsprinzipien als Basis für die Lösung von Randwertaufgaben der nichtlinearen Elastizitätstheorie," *ZAMM* Vol. 73, pp. T407–410.

Mang, H. A., Hofstetter, G., and Gallagher, R. H., 1985, "On an Alleged Philosopher's Stone Concerning Variational Principles With Side Conditions," *Int. J. Num. Meth. Engng.*, Vol. 21, pp. 1971–1999.

Xue, W.-M., and Atluri, S. N., 1985, "Existence and Stability, and Discrete BB and Rank Conditions, for General Mixed-Hybrid Finite Elements in Elasticity," R. Spilker, ed., ASME AMD-73, pp. 91–112.

APPENDIX A

Discussion of the Eqs. (40)

Integration of Eq. (36) gives

$$\tilde{\delta}\varphi = -\int y_1 dx \tilde{\delta}A - \int y_2 dx \tilde{\delta}B - \int y_p dx \tilde{\delta}C + \tilde{\delta}D, \quad \text{(A1)}$$

$$\tilde{\delta}w = -\int \left(\int y_1 dx\right) dx \tilde{\delta}A - \int \left(\int y_2 dx\right) dx \tilde{\delta}B - \int \left(\int y_p dx\right) dx \tilde{\delta}C + x\tilde{\delta}D + \tilde{\delta}E, \quad \text{(A2)}$$

where $\tilde{\delta}D$ and $\tilde{\delta}E$ are integration constants. (There is no danger of a confusion of the integration constant $\tilde{\delta}D$ in the Eqs. (A1) and (A2) with the integration constant $\tilde{\delta}D$ in Eq. (28). Hence, there is no need for different letter symbols for these constants.) Substitution of

$$\tilde{\delta}\epsilon = \frac{d}{dx}(\tilde{\delta}u) + \hat{\varphi}\tilde{\delta}\varphi, \quad \text{(A3)}$$

following from Eq. (5.2), into Eq. (23) yields

$$\frac{d}{dx}(\tilde{\delta}u) = \frac{1}{\mathsf{EA}}\tilde{\delta}C - \hat{\varphi}\tilde{\delta}\varphi. \quad \text{(A4)}$$

Integration of Eq. (A4) results in

$$\tilde{\delta}u = \int \frac{dx}{\mathsf{EA}}\tilde{\delta}C - \int \hat{\varphi}\tilde{\delta}\varphi dx + \tilde{\delta}F, \quad \text{(A5)}$$

where $\tilde{\delta}F$ is an integration constant.

It follows from the Eqs. (A2) and (A5) that

$$\tilde{\delta}u|_0^l = \int_0^l \frac{dx}{\mathsf{EA}}\tilde{\delta}C - \int_0^l \hat{\varphi}\tilde{\delta}\varphi dx, \quad \text{(A6)}$$

$$\tilde{\delta}w|_0^l = -\int_0^l \left(\int y_1 dx\right) dx \tilde{\delta}A - \int_0^l \left(\int y_2 dx\right) dx \tilde{\delta}B - \int_0^l \left(\int y_p dx\right) dx \tilde{\delta}C + l\tilde{\delta}D. \quad \text{(A7)}$$

With the help of the relations (A1), (A6), and (A7) the Eqs. (40) can be rewritten as

$$\begin{matrix}(40.1)\rightarrow\\(40.2)\rightarrow\\(40.3)\rightarrow\\(40.4)\rightarrow\end{matrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{Bmatrix} \tilde{\delta}A \\ \tilde{\delta}B \\ \tilde{\delta}C \\ \tilde{\delta}D \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \quad \text{(A8)}$$

where, e.g.,

$$a_{11} = -\left(\mathsf{E}Iy_1 \int y_1 dx\right)\Big|_0^l,$$

$$a_{12} = -\left(\mathsf{E}Iy_1 \int y_2 dx\right)\Big|_0^l, \quad a_{44} = l. \quad \text{(A9)}$$

The integration constants $\tilde{\delta}E$ and $\tilde{\delta}F$ do not depend on the solution of the above system of homogeneous linear algebraic equations. Hence, the SVP-based solution of the given problem

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

is only determined up to a rigid-body displacement of the beam in each one of the two coordinate directions.
If

$$\text{Det}\,[a_{ij}] \neq 0, \tag{A10}$$

then

$$\tilde{\delta}A = 0, \quad \tilde{\delta}B = 0, \quad \tilde{\delta}C = 0, \quad \tilde{\delta}D = 0. \tag{A11}$$

Substitution of the Eqs. (A11) into the Eqs. (A2) and (A5) yields

$$\tilde{\delta}u = \tilde{\delta}F, \quad \tilde{\delta}w = \tilde{\delta}E. \tag{A12}$$

Hence, the variations $\tilde{\delta}u$, $\tilde{\delta}w$, characterizing the nonuniqueness of the SVP-based solution of the given problem, are restricted to rigid-body displacements in the two coordinate directions.

This restriction will not hold, if

$$\text{Det}\,[a_{ij}] = 0. \tag{A13}$$

If

$$\text{Det}\begin{bmatrix} (\text{EI}y_1)_0 & (\text{EI}y_1)_l \\ (\text{EI}y_2)_0 & (\text{EI}y_2)_l \end{bmatrix} \neq 0, \tag{A14}$$

then, following from the Eqs. (40.1) and (40.2),

$$\tilde{\delta}\varphi_0 = 0, \quad \tilde{\delta}\varphi_l = 0. \tag{A15}$$

Substitution of the Eqs. (A15) into Eq. (40.3) yields

$$\tilde{\delta}u|_0^l = 0. \tag{A16}$$

In this case the eigenvalues obtained from Eq. (A13) must be equal to the ones obtained from

$$\text{Det}\,[b_{ij}] = 0, \tag{A17}$$

where $[b_{ij}]$ is the coefficient matrix of the following system of six homogeneous linear algebraic equations

$$\begin{matrix} \tilde{\delta}\varphi_l = 0 \rightarrow \\ \tilde{\delta}\varphi_0 = 0 \rightarrow \\ \tilde{\delta}u_l = 0 \rightarrow \\ \tilde{\delta}w_l = 0 \rightarrow \\ \tilde{\delta}u_0 = 0 \rightarrow \\ \tilde{\delta}w_0 = 0 \rightarrow \end{matrix} \begin{bmatrix} & \vdots & 0 & 0 \\ & \vdots & 0 & 0 \\ \bar{B} & \vdots & 0 & 1 \\ [6 \times 4] & \vdots & 1 & 0 \\ & \vdots & 0 & 1 \\ & \vdots & 1 & 0 \end{bmatrix} \begin{Bmatrix} \tilde{\delta}A \\ \tilde{\delta}B \\ \tilde{\delta}C \\ \tilde{\delta}D \\ \tilde{\delta}E \\ \tilde{\delta}F \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{A18}$$

where, e.g.,

$$b_{11} = \bar{b}_{11} = -\left(\int y_1 dx\right)_l, \quad b_{12} = \bar{b}_{12} = -\left(\int y_2 dx\right)_l,$$

$$b_{64} = \bar{b}_{64} = 0, \tag{A19}$$

with $\bar{b}_{11}$, $\bar{b}_{12}$, and $\bar{b}_{64}$ denoting coefficients of the $[6 \times 4]$-submatrix $\bar{B}$ of the above coefficient matrix. This assertion can be verified easily by subtracting the fifth row of the coefficient matrix in Eq. (A18) from the third row and the sixth row from the fourth row. The new third row represents the Eq. (A16) and the new fourth row the Eq. (40.4). Equation (A18) is based on the correct solution of the given problem, characterized by satisfaction of the field Eqs. (12.1) and (12.2) and, because of $\tilde{\delta}\varphi_l = 0$, $\tilde{\delta}\varphi_0 = 0$, ... $\tilde{\delta}w_0 = 0$, of the boundary conditions. Consequently, the eigenvalues obtained from Eq. (A13) are the correct eigenvalues. The eigenforms based on Eq. (A13) are characterized by an additional indeterminateness in comparison to the eigenforms based on Eq. (A17), namely, by indeterminate rigid-body displacements in the two coordinate directions. The eigenforms based on Eq. (A17) are the correct buckling modes.
If

$$\text{Det}\begin{bmatrix} (\text{EI}y_1)_0 & (\text{EI}y_1)_l \\ (\text{EI}y_2)_0 & (\text{EI}y_2)_l \end{bmatrix} = 0, \tag{A20}$$

then, following from the Eqs. (40.1) and (40.2),

$$\tilde{\delta}\varphi_l = \frac{(\text{EI}y_1)_0}{(\text{EI}y_1)_l}\tilde{\delta}\varphi_0 = \frac{(\text{EI}y_2)_0}{(\text{EI}y_2)_l}\tilde{\delta}\varphi_o. \tag{A21}$$

Consequently, in this case, the eigenforms based on Eq. (A13) do not satisfy the boundary conditions (A15). These eigenforms and the respective eigenvalues represent incorrect solutions of the given problem.

APPENDIX B

Discussion of the Eqs. (43)

Specialization of Eq. (28) for $\hat{C} = 0$ and integration of the resulting relations yields

$$\tilde{\delta}\kappa = -\frac{\hat{w}}{\text{EI}}\tilde{\delta}C - \frac{x}{\text{EI}}\tilde{\delta}D - \frac{1}{\text{EI}}\tilde{\delta}E, \tag{B1}$$

$$\tilde{\delta}\varphi = \int \frac{\hat{w}}{\text{EI}}dx\tilde{\delta}C + \int \frac{x}{\text{EI}}dx\tilde{\delta}D + \int \frac{1}{\text{EI}}dx\tilde{\delta}E + \tilde{\delta}G, \tag{B2}$$

$$\tilde{\delta}w = \int \left(\int \frac{\hat{w}}{\text{EI}}dx\right)dx\tilde{\delta}C + \int \left(\int \frac{x}{\text{EI}}dx\right)dx\tilde{\delta}D + \int \left(\int \frac{1}{\text{EI}}dx\right)dx\tilde{\delta}E + x\tilde{\delta}G + \tilde{\delta}H, \tag{B3}$$

where $\tilde{\delta}E$, $\tilde{\delta}G$, and $\tilde{\delta}H$ are integration constants. (There is no danger of a confusion of the integration constants $\tilde{\delta}D$ and $\tilde{\delta}E$ in the Eqs. (B1), (B2), and (B3) with the integration constants $\tilde{\delta}D$ and $\tilde{\delta}E$ in Eq. (A2). Hence, there is no need for different letter symbols for these constants.)

It follows from the Eqs. (B2) and (B3) that

$$\tilde{\delta}\varphi|_0^l = \int_0^l \frac{\hat{w}}{\text{EI}}dx\tilde{\delta}C + \int_0^l \frac{x}{\text{EI}}dx\tilde{\delta}D + \int_0^l \frac{1}{\text{EI}}dx\tilde{\delta}E, \tag{B4}$$

$$\tilde{\delta}w|_0^l = \int_0^l \left(\int \frac{\hat{w}}{\text{EI}}dx\right)dx\tilde{\delta}C + \int_0^l \left(\int \frac{x}{\text{EI}}dx\right)dx\tilde{\delta}D + \int_0^l \left(\int \frac{1}{\text{EI}}dx\right)dx\tilde{\delta}E + l\tilde{\delta}G. \tag{B5}$$

With the help of the relations (A6), (B4), and (B5) the Eqs. (43) can be rewritten as

$$\begin{matrix} (43.1) \rightarrow \\ (43.2) \rightarrow \\ (43.3) \rightarrow \end{matrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} & \vdots & 0 \\ c_{21} & c_{22} & c_{23} & \vdots & 0 \\ c_{31} & c_{32} & c_{33} & \vdots & 0 \end{bmatrix} \begin{Bmatrix} \tilde{\delta}C \\ \tilde{\delta}D \\ \tilde{\delta}E \\ \tilde{\delta}G \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{B6}$$

where

$$c_{11} = \int_0^l \frac{1}{\text{EA}}dx + \int_0^l \frac{\hat{w}^2}{\text{EI}}dx,$$

$$c_{12} = c_{21} = \int_0^l \frac{x\hat{w}}{\text{EI}}dx,$$

$$c_{13} = c_{31} = \int_0^l \frac{\hat{w}}{\text{EI}}dx, \; c_{22} = \int_0^l \frac{x^2}{\text{EI}}dx,$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$c_{23} = c_{32} = \int_0^l \frac{x}{EI}\,dx, \quad c_{33} = \int_0^l \frac{1}{EI}\,dx. \tag{B7}$$

It follows from the system of Eq. (B6) that $\tilde{\delta}G$ is indeterminate.

If

$$\mathrm{Det}[c_{ij}] \neq 0, \quad i, j = 1, 2, 3, \tag{B8}$$

then

$$\tilde{\delta}C = 0, \quad \tilde{\delta}D = 0, \quad \tilde{\delta}E = 0. \tag{B9}$$

Substitution of the Eqs. (B9) into the Eqs. (A5), (B2), and (B3) yields

$$\tilde{\delta}u = \tilde{\delta}F - \hat{w}\tilde{\delta}G, \quad \tilde{\delta}w = \tilde{\delta}H + x\tilde{\delta}G. \tag{B10}$$

Inserting Eq. (B10.1) and the relation

$$\tilde{\delta}\varphi = \tilde{\delta}G \tag{B11}$$

following from Eq. (B10.2) into Eq. (A3) gives

$$\tilde{\delta}\epsilon = 0. \tag{B12}$$

Hence, the variations $\tilde{\delta}u$, $\tilde{\delta}w$, characterizing the nonuniqueness of the SVP-based solution of the given problem, are restricted to a rigid-body motion consisting of rigid-body displacements $\tilde{\delta}F$ and $\tilde{\delta}H$ and of a rigid-body rotation $\tilde{\delta}G$. In contrast to the situation for $\hat{C} = 0$, for $\hat{C} \neq 0$ the variations $\tilde{\delta}u$, $\tilde{\delta}w$ are restricted to rigid-body displacements in the two coordinate directions, as has been shown previously.

The restriction of the variations $\tilde{\delta}u$, $\tilde{\delta}w$ to a rigid-body motion will not hold, if

$$\mathrm{Det}[c_{ij}] = 0, \quad i, j = 1, 2, 3. \tag{B13}$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

M. D. Gunzburger[2]
Interdisciplinary Center for
Applied Mathematics,
Virginia Tech.,
Blackburg, VA 24061-0531

H. C. Lee[3]
Department of Mathematics,
University of California,
Irvine, CA 92717

Feedback Control of Karman Vortex Shedding[1]

A computational study of the feedback control of the magnitude of the lift in flow around a cylinder is presented. The uncontrolled flow exhibits an unsymmetric Karman vortex street and a periodic lift coefficient. The size of the oscillations in the lift is reduced through an active feedback control system. The control used is the injection and suction of fluid through orifices on the cylinder; the amount of fluid injected or sucked is determined, through a simple feedback law, from pressure measurements at stations along the surface of the cylinder. The results of some computational experiments are given; these indicate that the simple feedback law used is effective in reducing the size of the oscillations in the lift.

1 Introduction

The control of the forces, e.g., lift and drag, exerted on a submerged obstacle by the fluid that surrounds it is important in many applications. Although here we focus on the problem of controlling the lift in the plane, unsteady, incompressible, viscous flow around a circular cylinder, the approach we take is applicable to other settings.

The control of flows has a long and rich history. Indeed, the control of flow about a cylinder was the subject of Prandtl's first paper! Some of these past efforts have been devoted to conducting experiments and simulations with different controlling mechanisms, without any attempt to optimally design or to actively change these mechanisms. For example, one may consult textbooks for detailed expositions of such efforts in the area of boundary layer control. Of course, there has been great success in the design and application of active controls, especially in the area of aerodynamic controls. These efforts, for the most part, use simple models for the effect of the flow on the submerged object.

Flows about a cylinder at even moderate values of Reynolds numbers, e.g., Re $>$ 47, exhibit an unsymmetric, periodic shedding of vortices. As a result, the lift force exerted by the fluid on the cylinder is periodic in time. Here, our goal is to show the efficacy of a simple feedback control law for the reduction of the magnitude of the lift. The context of our study is the full, nonlinear partial differential equations for plane, unsteady, incompressible, viscous flow. The control mechanism used to attempt to the reduce the size of the lift oscillations is the injection and suction of fluid through orifices on the surface of the cylinder. The amount of fluid injected or sucked through the orifices is determined, using a simple feedback law, from the pressure "measured" at various stations on the cylinder. Thus, in the language of feedback control, the *sensor* determines the pressure at various stations on the cylinder and the *actuator* injects or sucks fluid through orifices also on the cylinder. Other sensing and actuating mechanisms could also be used, e.g., the vorticity or tangential stress on the cylinder for sensing and rotation or shape modification of the cylinder for actuating.

Once a feedback law has been chosen, the computational simulation of the resulting flow is essentially no more difficult to accomplish than is the similar task for an uncontrolled problem. Of course, arbitrarily chosen feedback laws may not be effective in meeting the stated goal of reducing the size of the oscillations of the lift. However, there is evidence that appropriate controls can be applied to achieve a reduction in the amplitude of the oscillations of the lift. Small control mechanisms applied to the boundary layer can be effectively used to modify the generation of vortices and unsteady lift forces. This concept has been demonstrated by, e.g., Ffowcs and Zhao (1989).

The type of feedback laws used in our study are based on the observation that differences in the pressure on the top and bottom halves of the cylinder should give an indication of the asymmetric behavior of the lift. Such pressure differences are then used to determine how much fluid to inject or suck through the orifices. No attempt is made to systematically design an "optimal" feedback law, i.e., one that in some sense does the best possible job in meeting the objective. Future work will focus on the design of better feedback laws through, for example, the examination of systematic optimal control simulations. However, the computational results we have obtained indicate that the simple feedback law we use here is quite effective in reducing the size of the oscillations in the lift.

The plan of the rest of the paper is as follows. In the next section, details are given concerning the computational problem and algorithms that are used in the numerical simulations. Also, results are provided for the uncontrolled problem. Then, in Section 3, we report on some efforts to effect control using two symmetrically placed orifices on the back-side of the cylinder. In Section 4, results for three orifices are given which show that such an arrangement is useful in reducing the size of the oscillations in the lift.

2 The Uncontrolled Problem

2.1 The Computational Problem. Let B denote a circular cylinder and let Γ denote its boundary. The flow at infinity is assumed to be uniform and in the direction of the x-axis. We nondimensionalize with respect to the diameter of the cylinder and the velocity at infinity. Thus, in nondimensional variables, the cylinder has diameter unity and the flow at infinity has speed unity. The domain exterior to B is denoted by Ω. Let $\mathbf{u}(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$ denote the velocity and kinematic pressure (pressure divided by density), respectively. The governing equations, in

[1] This work was supported in part by the Air Force Office of Scientific Research under grant numbers AFOSR-93-1-0061 and AFOSR-F49620-93-1-0280 and by the Office of Naval Research under grant number N00014-91-J-1493.

[2] Current address: Department of Mathematics, Iowa State University, Ames, IA 50011-2064.

[3] Current address: Topology and Geometry Research Center and Department of Mathematics, Kyungpook National University, Taegu, Korea.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Mar. 10, 1995; final revision, June 27, 1995. Associate Technical Editor: T. R. Akylas.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

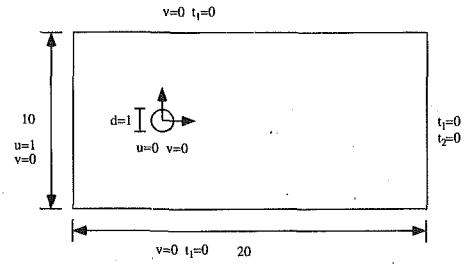


Fig. 1 The computational domain Ω for simulating flow past a circular cylinder

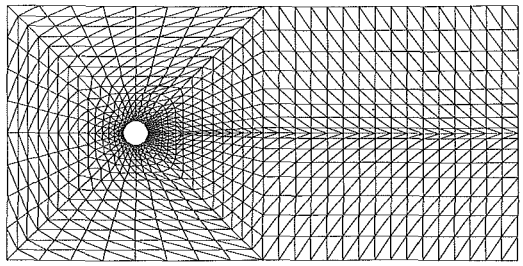


Fig. 2 The finite element grid for simulating flow past a circular cylinder

nondimensional form, for the unsteady, incompressible, viscous flow past the cylinder B are the continuity equation

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \text{ and for } \quad t > 0 \tag{1}$$

and the Navier-Stokes equations

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u} \quad \text{in } \Omega \text{ and for } \quad t > 0, \tag{2}$$

where Re denotes the Reynolds number based on the (dimensional) cylinder diameter d, the kinematic viscosity of the fluid ν, and speed at infinity U, i.e., $\mathrm{Re} = Ud/\nu$. Initial conditions are imposed on the velocity. Throughout, we start the flow impulsively from rest so that

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{0} \quad \text{in } \Omega. \tag{3}$$

On Γ, the surface of the cylinder, we prescribe the velocity. For the uncontrolled flow about the cylinder, the cylinder surface is a solid wall so that in this case

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma \text{ and for } \quad t > 0. \tag{4}$$

When controls are applied, this boundary condition becomes inhomogeneous on the portions of Γ covered by the injection and suction orifices. We defer discussion of such boundary conditions until feedback controls are introduced in Sections 3 and 4.

Since Ω is an exterior domain, one should also specify conditions that govern the behavior of the flow at large distances from the cylinder. We do not state these here since in the computational simulations a bounded flow domain will be used (see, e.g., Simo and Armero, 1993). Of course, one must choose an artificial computational domain and specify boundary conditions on the boundary of that domain. The combination of computational domain and computational boundary conditions should be chosen so that the resulting flow in the computational domain accurately approximates the flow in the exterior domain Ω.

The origin of the (x, y) coordinate system is located at the center of the cylinder and the cylinder has a unit diameter. The computational domain we use is the rectangle $-5 \le x \le 15$ and $-5 \le y \le 5$ with the cylinder excluded. The geometry of the domain is sketched in Fig. 1. The exterior boundary of the computational domain consists of the four sides of the rectangle; again, see Fig. 1. At the inflow boundary Γ_i the velocity is set to the uniform value at infinity, i.e., $u = 1$ and $v = 0$, where u and v denote the x and y components of the velocity vector, respectively. At the outflow Γ_o a vanishing "stress" boundary condition is imposed. Specifically, if $\mathbf{t} = -p\mathbf{n} + (1/\mathrm{Re})\partial \mathbf{u}/\partial n$, then on Γ_o we set $t_1 = t_2 = 0$, where t_1 and t_2 respectively denote the x and y components of $\mathbf{t}$. The vector $\mathbf{t}$ is not actually the stress vector; however, it has been found that imposing such an outflow condition is often more effective than requiring the true stress to vanish. Moreover, it is a more convenient boundary conditions to use with the particular form of the viscous term appearing in (2). On the top Γ_t and bottom Γ_b of the rectangle we impose the mixed conditions $v = 0$ and $t_1 = 0$. Summarizing the above discussion, on the exterior boundary of the computational domain we specify

$$u = 1 \quad \text{and} \quad v = 0 \quad \text{on } \Gamma_i \text{ and for } \quad t > 0, \tag{5}$$

$$t_1 = 0 \quad \text{and} \quad t_2 = 0 \quad \text{on } \Gamma_o \text{ and for } \quad t > 0, \tag{6}$$

and

$$v = 0 \quad \text{and} \quad t_1 = 0 \quad \text{on } \Gamma_b \text{ and } \Gamma_t \text{ and for } \quad t > 0. \tag{7}$$

The boundary conditions are also indicated on the sketch of the computational domain given in Fig. 1.

For the uncontrolled case, the problem we wish to discretize is then defined by (1)–(7), where in (1)–(3) Ω now denotes the computational domain sketched in Fig. 1.

The spatial discretization is effected using the Taylor-Hood finite element pair based on a triangulation of the computational flow domain (see, e.g., Girault and Raviart, 1986; Gunzburger, 1989; Taylor and Hood, 1973). A typical triangulation is depicted in Fig. 2; it consists of 1735 triangles and 3735 (velocity) nodes. The velocity field is approximated using continuous piecewise quadratic isoparametric functions; the pressure field is approximated by continuous piecewise linear isoparametric functions; the same grid is used for both fields; see Fig. 3.

The semi-implicit single-step Euler method is employed for the time discretization. This time-stepping algorithm is defined as follows: given $\mathbf{u}^{m-1}$, the pair $(\mathbf{u}^m, p^m)$ is obtained by solving the linear system

$$\nabla \cdot \mathbf{u}^m = 0 \tag{8}$$

and

$$\frac{\mathbf{u}^m - \mathbf{u}^{m-1}}{\delta} + (\mathbf{u}^{m-1} \cdot \nabla)\mathbf{u}^m + \nabla p^m = \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u}^m, \tag{9}$$

where δ denotes the time step. Of course, (8) and (9) are supplemented by the appropriate boundary and initial conditions. This scheme has been analyzed in Temam (1984).

To break the symmetry in the problem and trigger the vortex shedding, a perturbation is applied for a short time interval. A geometric imperfection may also be applied for this purpose

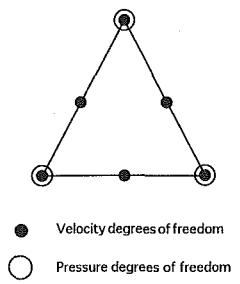


Fig. 3 The Taylor-Hood element pair

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

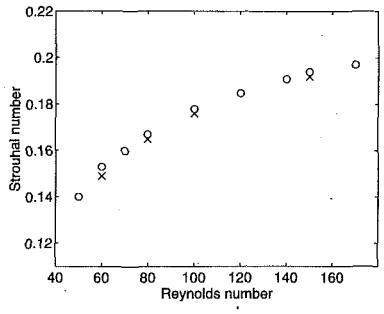


Fig. 4 Plot of the relationship between Strouhal and Reynolds numbers for the natural shedding respose. "x": our numerical results, "o": the numerical results reported by Karniadakis et al.

(see, e.g., Simo and Armero, 1993). The transient solution develops until a periodic state is reached.

2.2 Results for the Uncontrolled Problem. Using the computational scheme described above, we have simulated the uncontrolled problem at the Reynolds numbers of 60, 80, 100, and 150. The Strouhal number predicted by our numerical calculation is higher than the experimental ones (Roshko, 1954; Williamson, 1989) by about six to eight percent. The disagreement between our numerical computations and experiments (Roshko, 1954; Williamson, 1989) can be partly attributed to the truncation of the computational domain in our numerical simulations and also to the experimental measurements (see Karniadakis and Triantafyllou, 1989). Our computed value of the Strouhal number is in agreement with the one computed by Gresho et al. (1984) using a modified finite element method and the one computed by Karniadakis and Triantafyllou (1989) using a spectral element method. The nondimensional frequency St is defined as $f_d d/U$, where f_d is the frequency of vortex shedding. In Fig. 4, our results are plotted along with the one computed by Karniadakis and Triantafyllou (1989).

The periodic state is reached at approximately $t = 180$ and $t = 130$ for Re = 60 and Re = 80, respectively. These phases of the solution are perfectly captured by the lift coefficient C_L and the drag coefficient C_D defined by

$$C_L = \int_0^{2\pi} \tilde{t}_2(\theta)d\theta, \tag{10}$$

and

$$C_D = \int_0^{2\pi} \tilde{t}_1(\theta)d\theta, \tag{11}$$

respectively, where $\tilde{t}_2$ is the y component of the true stress vector, $\tilde{t}_1$ is the x component of the true stress vector, and θ is the angle along the surface of the surface of the cylinder measured from the leading edge. In Figs. 5 and 6, we give the evolutions in time of C_L and C_D for Re = 60 and Re = 80, respectively. The evolutions in time of C_L shown in Figs. 5 and 6 exhibit the oscillations having a period $T \approx 6.7$ units in time for Re = 60 and $T \approx 6.0$ for Re = 80, leading to a Strouhal number $St \approx 0.149$ for Re = 60 and $St \approx 0.165$ for Re = 80.

We have already mentioned that we will sense the pressure at stations on the surface of the cylinder (see Fig. 7). It is therefore instructive to examine the oscillatory behavior of that pressure distribution. In Fig. 8 is given the pressure distribution along the surface of the cylinder at different times during one period of oscillation of the lift coefficient, i.e., in a time interval between successive sheddings of vortices from the same location on the cylinder. A zero angle corresponds to the forwardmost point on the cylinder. As is expected, the pressure distribution on the cylinder is asymmetric. This is especially evident at the angles $\pm\pi/2$, i.e., at the top and bottom of the cylinder.

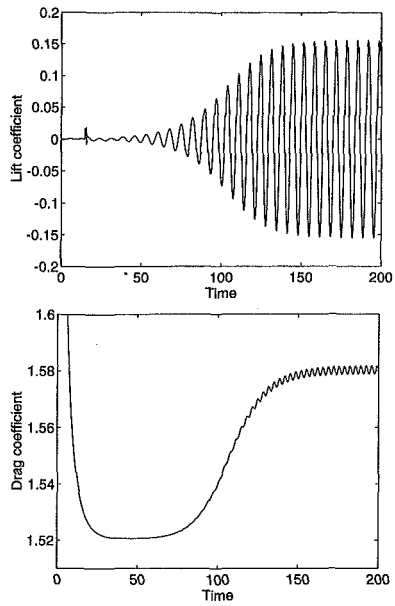


Fig. 5 The evolution in time of lift coefficient C_L and drag coefficient C_D for Re = 60

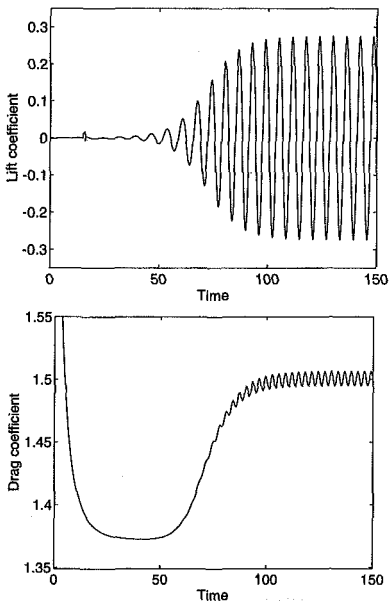


Fig. 6 The evolution in time of lift coefficient C_L and drag coefficient C_D for Re = 80

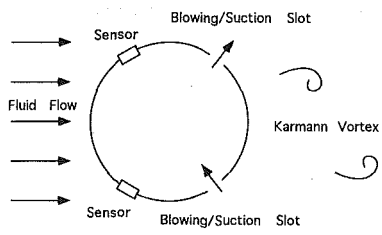


Fig. 7 Locations of sensors and actuators

The oscillatory nature of the pressure is also evident from Fig. 9 in which the pressure at some specific locations on the cylinder is graphed versus time. Note that the period of the

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

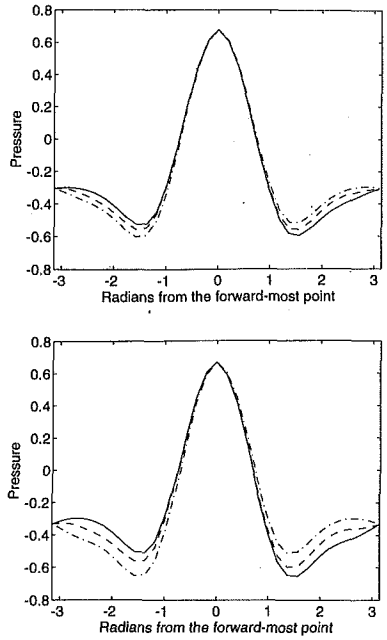


Fig. 8 Pressure distributions along the cylinder surface at different times for Re = 60 (top) and Re = 80 (bottom)

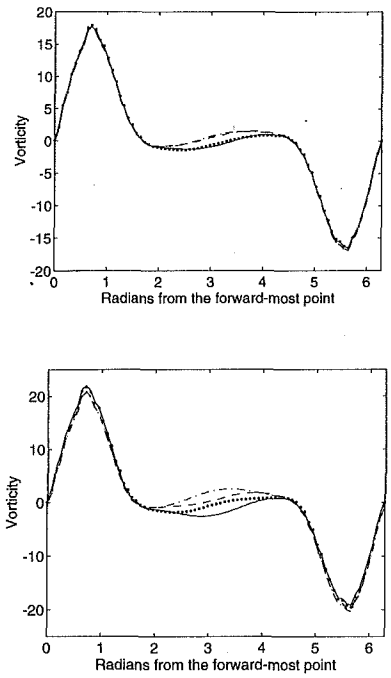


Fig. 10 Vorticity distributions along the cylinder surface at different times for Re = 60 (top) and Re = 80 (bottom)

oscillations are independent of the location, although the magnitudes are not.

Instead of the pressure, one could also sense the vorticity on the surface of the cylinder. For a symmetric flow, i.e., one where u and v are even and odd functions of y, respectively, the pressure and vorticity are even and odd function of y, respectively. As can be inferred from Fig. 10, the vorticity distribution along the surface of the cylinder deviates from that of a symmetric flow, although the deviations are less pronounced than those of the pressure. In this case, the region around $\pm\pi$ is a good candidate for the sensor location since the maximum deviation of vorticity occurs at point $\pm\pi$.

3 Control by a Pair of Suction/Blowing Orifices

Perhaps the most "natural" arrangement of orifices is a pair of blowing/suction slots located symmetrically on the backside of the cylinder. We have tried such arrangements with slots of length $\pi/16$ centered about the angles $\pm 11\pi/32$, $\pm 15\pi/32$, or $\pm 19\pi/32$ from the leading edge of the cylinder. The separation points on the cylinder surface at a Reynolds number of 60 wander in the vicinity of $\pm 2\pi/3$ as the vortices are shed. When the blowing/suction slots are located slightly ahead of the separation points of the uncontrolled wake, they proved to be a strong actuator which is in agreement with Park et al. (1993).

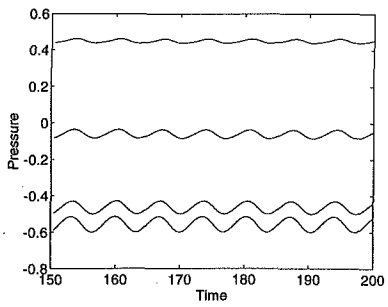


Fig. 9 The evolution in time of pressure at $\theta = \pi/8$, $\pi/4$, $3\pi/8$, and $\pi/2$ (from top to bottom) for Re = 60

The vortex shedding behind a circular cylinder can be suppressed by using the following feedback control law:

$$\mathbf{u} = \alpha \frac{\omega(t, x_s, y = 0)}{\omega_{\max}} \mathbf{g}(\theta), \tag{12}$$

where α is the feedback coefficient, $\omega(t, x_s, y = 0)$ is the measured vorticity at the downstream location $(x_s, 0)$ along the centerline, and $\omega_{\max}$ is the maximum vorticity at the same spatial location for the unforced wake. Thus, sensing is done at a point downstream of the cylinder. Here, we found two major problems. First, successful control is possible for only a very narrow range of phases while the vortex shedding is enhanced if forced with a phase out of this range. Second, the "optimal" location of sensor for Re = 60 is not an optimal location for Re $\neq$ 60. (Of course, placing sensors in the middle of the flow is problematic in itself.)

We try some more robust controls to overcome these problems. We place the pressure sensors at symmetric locations on the front side of the cylinder. From Figs. 8 and 9, one may conclude that the specific position of the sensors is not overly important; we have found that to be true. For the results given below, we placed the sensors at $\pm 7\pi/32$ from the leading edge.

We first determine the blowing/suction velocity through the orifices as dictated by the following feedback control law:

$$\mathbf{u} \cdot \mathbf{n} = \operatorname{sign}\{p(\theta_1) - p(\theta_2)\} \times \min\{\alpha | p(\theta_1) - p(\theta_2)|, \beta\} g(\theta) \tag{13}$$

and the tangential velocity vanishing. Here $g(\theta)$ is a given function, θ_1 and θ_2 are the positions of the sensors on the cylinder, α is the feedback coefficient, and $\beta > 0$ is a parameter that limits the size of the control. The function g can be used to specify the flow profile at the orifices. Here g is a parabolic (quadratic) function along the two orifices with maximum value 1 and minimum value -1, respectively. Our computational results show that for any value of α the system quickly becomes unstable, i.e., at every time step each orifice switches from blowing to suction and back to blowing, etc. We found that the pressure on the surface of the circular cylinder is very sensitive to a velocity perturbation at the actuators.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

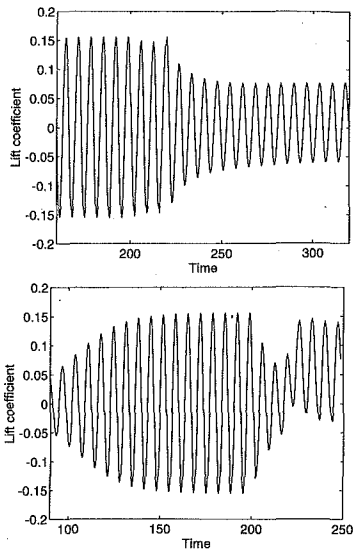


Fig. 11 The evolution of the lift coefficient for case (*a*) (top) and case (*b*) (bottom) where Re = 60

Next, we try another approach called *digital control* which is based on the *sample-and-hold* technique. The values $p(\theta_1)$ and $p(\theta_2)$ in (13) are measured only at discrete instants or *sampling times.* The control law is updated at each time $k\Delta t$, k an integer, on the basis of the sampled values $p(\theta_1)$ and $p(\theta_2)$. According to our experience, a choice of Δt in the range $2\delta \leq \Delta t \leq 5\delta$ resulted in a stable configuration. If we set $\Delta t = 5\delta$, $\alpha = 3$, and $\beta = 0.1$, we found that the magnitude of the oscillations in the lift coefficient was reduced by a factor of one-half. If $\alpha = 4$, the lift coefficient was shifted upwards which implies that one orifice stays in a blowing state for long periods of time and in a sucking state for only a short period. If $\beta > 0.1$, the Karman vortex shedding was enhanced. In Figs. 11 and 12, the results of two particular computations are given. These correspond to the two cases:

(a) $\alpha = 3.0$,
(b) $\alpha = 4.0$,

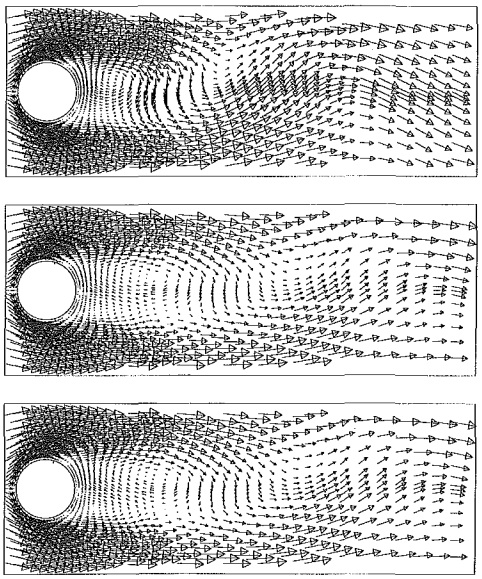


Fig. 12 The velocity field for the uncontrolled case, case (*a*), and case (*b*) (from top to bottom) at nondimensional time 299.5 where Re = 60

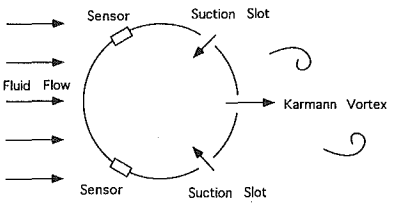


Fig. 13 Locations of sensors and actuators

with $\beta = 0.1$ for both cases. The control was turned on at $t = 200$. In Fig. 12, we can see that the velocity fields for the controlled cases is more symmetric than that for the uncontrolled case.

We have also used this control strategy for Re = 70 and 80. In the case Re = 70, we obtained similar results to those discussed above for Re = 60. For Re = 80, we have found that we can not suppress the vortex shedding behind circular cylinder using the above control mechanism. Thus, we next tried another control mechanism.

4 A Pair of Suction Slots With a Single Blowing Slot

To the pair of orifices located at $\pm 23/32\pi$ from the leading edge, we add an orifices at $\theta = \pi$, i.e., at the trailing edge of the cylinder. The sizes of all three orifices are same (size $\pi/32$). Specifically, fluid is sucked through two slots centered at $\pm 23\pi/32$ and is blown through the slot centered at π (see Fig. 13). We again sense the pressure at two locations θ_1 and θ_2 on the front face of the cylinder. The position of the sensors and actuators remained fixed for computations with different values of the Reynolds number. The feedback control law is chosen as follows: at the ith orifice we have

$$\mathbf{u}_i \cdot \mathbf{n} = \min \{\alpha | p(\theta_1) - p(\theta_2)|, \beta\} g_i(\theta) \qquad (14)$$

and the tangential velocity vanishing. Here, again, g_i is a parabolic function along the orifices.

We have performed computational simulations for numerous cases. Since we have obtained similar results for various cases, we report only the following cases:

(c) $\alpha = 40$, $\beta = 2.0$,
(d) $\alpha = 80$, $\beta = 2.0$,
(e) $\alpha = 80$, $\beta = 1.5$.

In our computation, we have that on the suction orifices the maximum value of $|g_i|$ is 0.5, while on the single blowing orifice the maximum value is 1. Thus, we have that $\int_{\partial B} \mathbf{u} \cdot \mathbf{n} dS = 0$ so that the amount of fluid blown through the slot located on the x-axis was twice the amount of fluid sucked through each of the two off-axis slots, i.e., there is no net fluid introduced into the flow. In all cases, the control was turned on at $T = 140$, a value of the nondimensionalized time at which the Karman vortex street pattern has fully developed.

In Fig. 14, we give the evolution of the control, the lift coefficient and drag coefficient for the case (c). A sudden change of boundary condition causes a very stiff drag coefficient in time. After some transition time, this flow reaches another

Table 1 Summary of results for cases (c), (d), and (e)

Results case	Control coefficient, C_C (average)	Lift coefficient, C_L (peak-to-peak)	Drag coefficient, C_C (average)
No control	0.0	0.5489	1.5006
Case (c)	0.8931	0.0875	1.4878
Case (d)	0.9521	0.0589	1.4865
Case (e)	0.9628	0.0634	1.4951

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

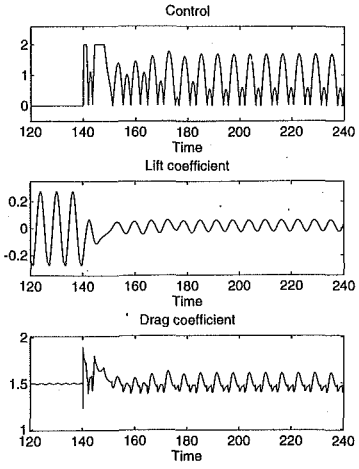


Fig. 14 The evolution of the control, the lift coefficient and drag coefficient (from top to bottom) for case (*c*) where Re = 80

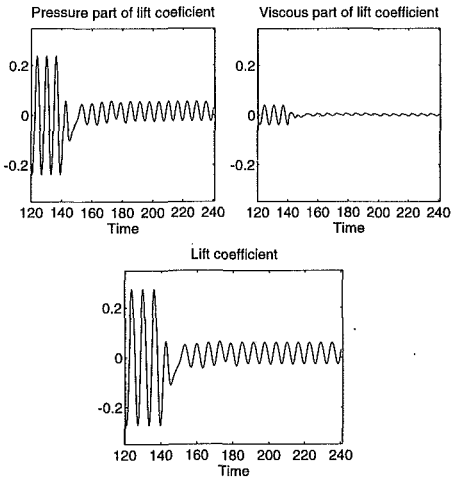


Fig. 15 The evolution of the lift coefficient (bottom), the pressure contribution of the lift coefficient (top left) and the viscous contribution of the lift coefficient (top right) for case (*c*) where Re = 80

periodic solution. The average control and drag coefficient are 0.8931 and 1.4878, respectively. The average control and drag coefficient are defined by

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T C_C(t)dt \quad \text{and} \quad \lim_{T\to\infty}\frac{1}{T}\int_0^T C_D(t)dt, \qquad (15)$$

respectively, where $C_C = \min\{\alpha|p(\theta_1) - p(\theta_2)|, \beta\}$. The oscillations in the lift are reduced to approximately 16 percent of that for the uncontrolled case. In Fig. 15, we give the pressure and viscous contributions of the lift coefficient.

In Fig. 16, we give the evolution of the control, the lift coefficient and drag coefficient for the case (d). Note, from the results for the cases (c) and (d), that increasing the value of α results in smaller oscillations since the available control can be larger. In this case, the flow reaches a quite chaotic solutions after some transition time. The average control and drag coefficient are 0.9521 and 1.4865, respectively. The oscillations in the lift are reduced to approximately 11 percent of that for the uncontrolled case. In Fig. 17, we give the pressure and viscous contributions of the lift coefficient.

In case (e), instead changing α we change the value of β. In Fig. 18, we give the evolution of the control, the lift coefficient and drag coefficient for the case (e). The average control and drag coefficient are 0.9628 and 1.4951, respectively. The

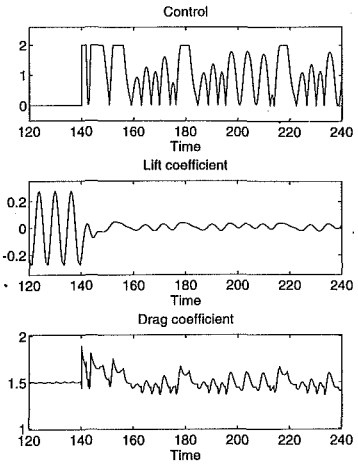


Fig. 16 The evolution of the control, the lift coefficient and drag coefficient (from top to bottom) for case (*d*) where Re = 80

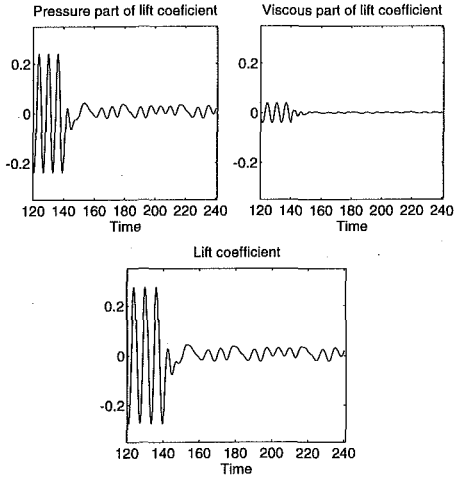


Fig. 17 The evolution of the lift coefficient (bottom), the pressure contribution of the lift coefficient (top left) and the viscous contribution of the lift coefficient (top right) for case (*d*) where Re = 80

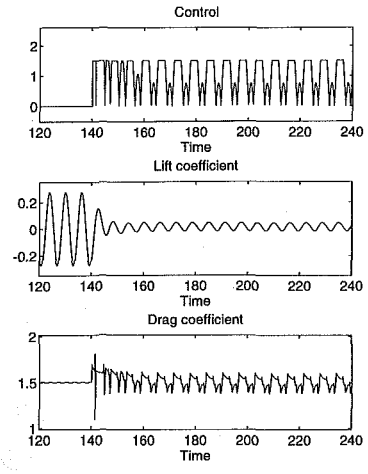


Fig. 18 The evolution of the control, the lift coefficient and drag coefficient (from top to bottom) for case (*e*) where Re = 80

oscillations in the lift are reduced to approximately 12 percent of that for the uncontrolled case. From the results for case (e), one notes that reducing the value of β can result in a more

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

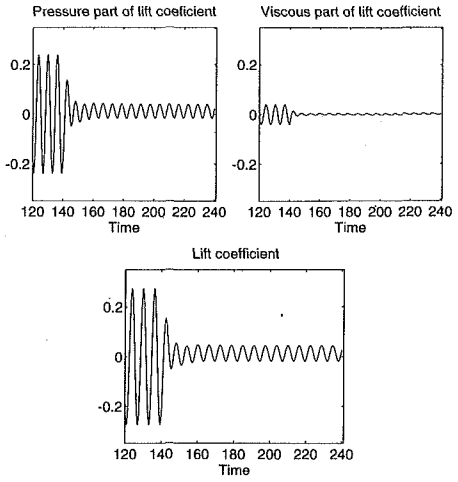


Fig. 19 The evolution of the lift coefficient (bottom), the pressure contribution of the lift coefficient (top left) and the viscous contribution of the lift coefficient (top right) for case (*e*) where Re = 80

regular, i.e., periodic, behavior for the reduced lift coefficient like the results in the case (c). In Fig. 19, we give the pressure and viscous contributions of the lift coefficient.

Comparing the case (e) with the case (c), one notes that the oscillation in the lift for the case (e) is smaller than that for the case (c) while the drag coefficient for the case (e) is larger than that for the case (c). From the results of the cases (d) and (e), one also can see that the oscillation in the lift can not be reduced while the average control is increased. From Figs. 15, 17, and 19, we see that the oscillations in the lift are reduced by mainly the change of the flow pattern, i.e., the unsymmetric flow is changed to a symmetric flow. We also see that the simple feedback law (14) is quite effective in reducing the size of the oscillations in the lift coefficient without increasing the drag coefficients.

In Fig. 20, the velocity field is given for the uncontrolled case and for the cases $(c)-(e)$ at $t = 340.05$. In Fig. 21, the horizontal velocity fields for the uncontrolled case and for case (d) are given.

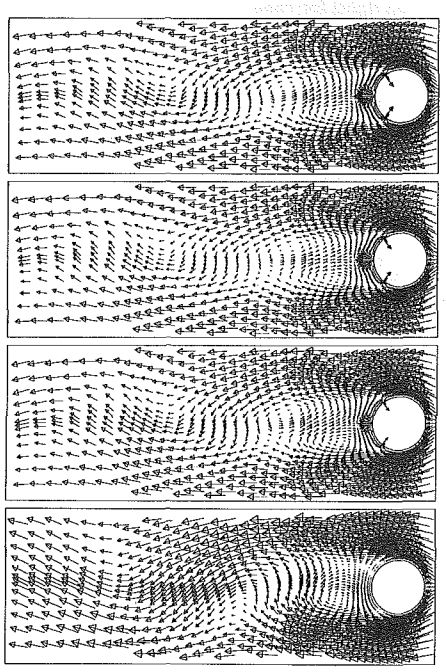


Fig. 20 The velocity field for the uncontrolled case, case (*c*), case (*d*), and case (*e*) (from top to bottom) at $t = 340.05$ where Re = 80

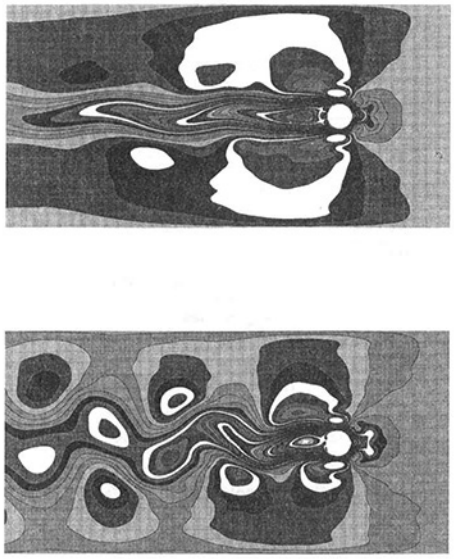


Fig. 21 The horizontal velocity for uncontrolled (top) and controlled (bottom, case (*d*)) flows around cylinder at $t = 340.05$ where Re = 80

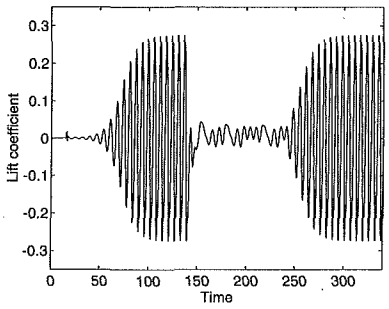


Fig. 22 The evolution of the lift coefficient for case (*d*) where Re = 80

In Fig. 22 we see what happens of if we turn the control off. Specifically, we had the parameters set as in case (d); the control (forcing) was turned on at $t = 140$ and turned off at $t = 240$. We see that, before the control is turned on, we have large oscillations in the lift coefficient. After the control is turned on at $t = 140$, the size of the oscillations are greatly reduced. However, after the control is turned off at $t = 240$, the size of the oscillation in lift coefficient once again become relatively large; in fact, the flow returns to the Karman vortex street structure present before the control was turned on.

5 Conclusions

By adopting a simple arrangement of pressure sensors and injection/suction actuators and a simple feedback law, we have been able to reduce the size of the oscillations in the lift quite effectively without increasing the drag coefficient. These feedback laws are not systematically designed and are not "optimal" feedback laws. The optimal feedback control of Karman vortex shedding is currently being studied.

References

Beger, E., and Wille, R., 1972, "Periodic Flow Phenomena," *Ann. Rev. Fluid Mech.,* Vol. 4, pp. 313–340.

Ffowcs, J. E., Williams, and Zhao, B. C., 1989, "Active Control of Vortex Shedding," *J. of Fluids and Structure,* Vol. S3, pp. 115–122.

Girault, V., and Raviart, P. A., 1986, *Finite element methods for Navier-Stokes equations,* Springer, Berlin.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Gresho, P., Chan, S., Lee, R., and Upson, C., 1984, "A Modified Finite Element Method for Solving the Time-Dependent Incompressible Navier-Stokes Equations, Part II: Applications," *Intl J. Num. Mech. Fluids,* Vol. 4, pp. 619–640.

Gunzburger, M., 1989, *Finite element methods for incompressible viscous flows: A guide to theory, practice and algorithms* Academic, Boston, MA.

Karniadakis, G., and Triantafyllou, G., 1989, "Frequency Selection and Asymptotic States in Laminar Wakes," *J. Fluid Mech.,* Vol. 199, pp. 441–469.

Oertel, H., Jr., 1990, "Wakes Behind Blunt Bodies," *Annu. Rev. Fluid Mechanics,* Vol. 22, pp. 539–564.

Park, D. S., Ladd, D. M., and Hendricks, E. W., 1993, "Feedback Control of Karman Vortex Shedding Behind a Circular Cylinder at Low Reynolds Numbers," *Phys. Fluids A.,* submitted for publication.

Roshko, A., 1954, "On the Development of Tubulent Wakes from Vortex Streets," *Natl Advisory Committee Aeronaut., Rep,* Vol. 1191.

Simo, J. C., and Armero, F., 1993, "Unconditional Stability and Long-Term Behavior of Transient Algorithms for the Incompressible Navier-Stokes and Euler Equations," Stanford University, Stanford, CA.

Taylor, C., and Hood, P., 1973, "A Numerical Solution of the Navier-Stokes Equations Using the Finite Element Methods," *Comput. & Fluids,* Vol. 1, pp. 73–100.

Temam, R., 1984, *Navier-Stokes Equations,* North-Holland, Amsterdam.

Williamson, C. H. K., 1989, "Defining a Universal and Continuous Strouhal-Reynolds Number Relationship for the Laminal Vortex Shedding of a Circular Cylinder," *J. Fluid Mech.,* Vol. 206, pp. 579–627.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Cavity-Opening Tensor Method and its Application to Reflection and Transmission of Ultrasonic Wave by Defects

Yonglin Xu
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208

The reflection and transmission of ultrasonic waves by an array of planar defects are investigated using the cavity-opening tensor (COT) method, whereby each defect or flaw can be simulated as a three-dimensional tensor in terms of the displacement distribution over the surface of the defect. The closed-form expressions of reflection and transmission coefficients represented by one set of COTs are derived, and the interaction factors of COTs are computed using the boundary integral equation with the approximate integral only on the single defect surface. Moreover, the analytical expressions of scattered fields at a distance from the defects are given by the representation of COT. As compared with the conventional numerical method, the COT method greatly simplifies the calculations, and also makes it possible to more efficiently characterize defects with more complicated distributions.

1 Introduction

Diffusion bonds may have three kinds of defects: Volumetric defects (voids or inclusions), crack-like defects, and crystalline grains. These defects, generally distributed in the plane of the bond, provide a plane of partial reflection and transmission for ultrasonic wave motion, and hence the ultrasonic method can, in principle, be used to detect and characterize bond-plane defects.

The boundary element method has been widely used in the numerical analysis of reflected and transmitted waves for an array of defects. For a diffusion bond with a great number of defects, the large number of elements makes such numerical analysis somewhat cumbersome. Therefore, the cavity-opening tensor (COT) method is developed for simulating each defect or cavity as a three-dimensional tensor in terms of the displacement over the surface of the defect, and the scattered field and coefficients of the reflection and transmission can be estimated in terms of the values of the COTs. This simulation ultimately provides information on the bond strength of the plane bonds. For the special case of crack problems, the tensor can be reduced to a vector or a scalar.

Based on an assumption that the interactions of the neighboring two cavities are approximately linear, the interaction factors of cavities can be estimated using the boundary integral equations with the integral only on a single defect surface. The calculations show that the interaction factors are inversely proportional to the distance between two cavities, and therefore the nearest cavities have an appreciable influence on the COT. For the special case of periodic distributions, the reflection and transmission coefficients computed as functions of the frequency agree very well with the other rigorous numerical results.

Reflection and transmission of time-harmonic waves by an array of cracks, cavities, and inclusions have been intensively studied for nondestructive evaluation of defects in solids (Achenbach and Kitahara, 1986, 1987; Angel and Achenbach, 1985, 1987; Mikata and Achenbach, 1988; Achenbach and Li, 1986; Mikata 1993). Reflection and transmission coefficients for an array of cavities have been obtained by Achenbach and Kitahara (1986) using the boundary integral equation method. Reflection and transmission of elastic waves by an array of coplanar cracks were studied by Angel and Achenbach (1985, 1987) using the Fourier series expansions of the Green-Lame potential, and problems of a periodic array of inclined cracks and randomly oriented cracks were solved by Mitaka (1988a, b). Wave propagation in a periodically layered, infinite elastic body was investigated by Delph et al. (1978, 1979), and the dispersion equation is obtained by using Eloquet's theory. Diffusion bonds between two pieces of identical material were studied by Baik and Thompson (1984) and Margetan et al. (1992) based on the use of a distributed spring model for determining the bond strength and other mechanical properties.

For nonperiodic distributions, arrays of cracks by normal incident waves were considered by Sotiropoulos and Achenbach (1988) using the crack-opening volume (COV) method, in which a crack is simulated as a scalar, the crack-opening volume, in terms of the crack-opening displacement. The COV method can simplify the calculation of the reflection and transmission for an array of cracks. However, applications of the COV method are limited to the case of cracks by normal incident waves. Thus, the COV method has been extended to a more efficient method, the COT method, for the analysis of ultrasonic waves and the detection of defects with more complicated geometric configurations.

2 Cavity-Opening Tensor Method for Scattering

Although the scattering of a plane longitudinal wave by cavities can be solved by the integral equation method, the COT method may provide a more effective approach to the analysis of scattered fields. As shown in Fig. 1, a plane longitudinal time-harmonic wave, $\mathbf{u}^I \exp(-i\omega t)$, is incident on a single cavity, where $\mathbf{u}^I$ is given by

$$\mathbf{u}^I = \mathbf{p} \exp(iK_L \mathbf{p} \cdot \mathbf{x}) \qquad (2.1a)$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Manuscript received by the ASME Applied Mechanics Division, Apr. 12, 1995; final revision, Oct. 22, 1995. Associate Technical Editor: D. M. Barnett.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

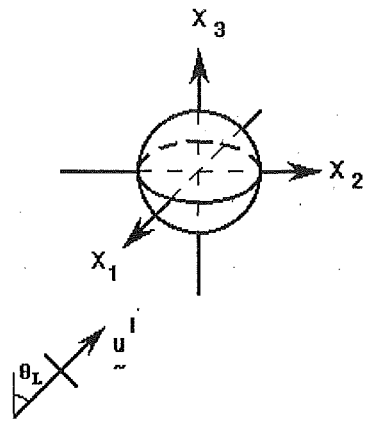


Fig. 1 The longitudinal incident wave is incident on a cavity

Here, $\mathbf{p}$, the unit vector with the angle of incidence θ_L, is written as

$$\mathbf{p} = (0, \sin\theta_L, \cos\theta_L) \tag{2.1b}$$

and

$$K_L = \omega/C_L, \quad C_L = \sqrt{\frac{\lambda + 2\mu}{\rho}} \tag{2.2}$$

where the Lame constants λ, μ and the density ρ characterize the elastic solid. The time-harmonic term $\exp(-i\omega t)$, here ω is the angular frequency, will be omitted in the sequel. The surfaces of the cavities are free of tractions,

$$t_i = \tau_{ij} n_j = 0 \tag{2.3}$$

where n_j is the component of the unit vector to the cavity surface S, and σ_{ij} denotes the stress tensor. The scattered field at a distance from the defect can be determined by the value of the cavity-opening tensor. Based on the Betti-Rayleigh reciprocity principle, the displacement induced by a plane longitudinal time-harmonic wave can be written as

$$u_i(\mathbf{x}) + \sum_{n=1}^{N} \int_{s_n} T_{ij}^{*}(\mathbf{x}, \mathbf{y}) u_j(\mathbf{y}) ds(\mathbf{y}) = u_i^I(\mathbf{x}) \tag{2.4}$$

and thus, the scattered field u_i^s satisfies the integral equation

$$u_i^s(\mathbf{x}) = -\sum_{n=1}^{N} \int_{s_n} T_{ij}^{*}(\mathbf{x}, \mathbf{y}) u_j(\mathbf{y}) ds(\mathbf{y}) \tag{2.5}$$

where the scattered field is a superposition of N integrals which are determined by the displacements over surfaces of the cavities. For one of the cavities as shown in Fig. 2, the Green function can, in terms of $T_{ij}^{*} = \sigma_{ijk}^{*} n_k$, be written as

$$T_{ij}^{*} = (A_{ijk}^{L} e^{ik_L r} + A_{ijk}^{T} e^{ik_T r}) n_k \tag{2.6}$$

where functions A_{ijk}^{L} and A_{ijk}^{T} are

$$A_{ijk}^{L} = \frac{1}{4\pi\mu}\left\{\lambda\delta_{kj}\left(\frac{K_L}{K_T}\right)^2 r_{,i} a_{1L} - \frac{2\mu}{K_T^2}[(\delta_{ij} r_{,k} + \delta_{jk} r_{,i} + \delta_{ki} r_{,j}) a_{2L} + r_{,i} r_{,j} r_{,k} r^2 a_{3L}]\right\}$$

$$A_{ijk}^{T} = \frac{1}{4\pi\mu}\left\{\mu(\delta_{ij} r_{,k} + \delta_{ik} r_{,j}) a_{1T} + \frac{2\mu}{K_T^2}[(\delta_{ij} r_{,k} + \delta_{jk} r_{,i} + \delta_{ki} r_{,j}) a_{2T} + r_{,i} r_{,j} r_{,k} r^2 a_{3T}]\right\}. \tag{2.7a, b}$$

In Eqs. (2.7a, b), $r^2 = \Sigma (x_i - y_i)^2$ and $r_{,j} = \partial r/\partial y_j = -(\partial r/\partial x_j)$. The coefficients $a_{j\alpha}$ ($j = 1, 2, 3$; $\alpha = L, T$) are

$$a_{1\alpha} = \frac{1}{r}\left(iK_\alpha - \frac{1}{r}\right).$$

$$a_{2\alpha} = -\frac{1}{r^2}\left(K_\alpha^2 + 3iK_\alpha \frac{1}{r} - 3\frac{1}{r^2}\right)$$

$$a_{3\alpha} = -\frac{1}{r^3}\left(iK_\alpha^3 - 6K_\alpha^2 \frac{1}{r} - 15iK_\alpha \frac{1}{r^2} + 15\frac{1}{r^3}\right)$$

$$(\alpha = L, T). \tag{2.8a, b, c}$$

For details of the derivation one can refer to Xu (1992). When the distance between the observation point and the cavity center becomes larger (Fig. 2), it can be written as

$$|\mathbf{x}| = |\mathbf{r}| - \mathbf{r}_0 \cdot \mathbf{R}_0 \tag{2.9}$$

where $\mathbf{r}_0 = \mathbf{r}/|\mathbf{r}|$ and $\mathbf{R}_0$ is a radius vector. The use of (2.5), (2.6), and (2.9) gives the scattered field for a single cavity as

$$u_i^s(\mathbf{x}) = -A_{ijk}^{L}(\mathbf{r}) e^{ik_L r} \int_s u_j n_k e^{-ik_L \mathbf{r}_0 \cdot \mathbf{R}_0} ds(\mathbf{y}) - A_{ijk}^{T}(\mathbf{r}) e^{ik_T r} \int_s u_j n_k e^{-ik_T \mathbf{r}_0 \cdot \mathbf{R}_0} ds(\mathbf{y}) \tag{2.10}$$

where the term of the surface integral acts as a source at the cavity center and stands by V_{jk}^{α}, which is defined as a cavity-opening tensor. Thus, the scattered field can be expressed as

$$u_i^s(\mathbf{r}) = -A_{ijk}^{L}(\mathbf{r}) e^{ik_L r} V_{jk}^{L} - A_{ijk}^{T}(\mathbf{r}) e^{ik_T r} V_{jk}^{T} \quad (i = 1, 2, 3) \tag{2.11}$$

where the cavity-opening tensor is

$$V_{jk}^{\alpha} = \int_s u_j n_k e^{-ik_\alpha \mathbf{r}_0 \cdot \mathbf{R}_0} ds(\mathbf{y}), \quad (\alpha = L, T). \tag{2.12}$$

The functions $A_{ijk}^{L}(\mathbf{r}) e^{ik_L r}$ and $A_{ijk}^{T}(\mathbf{r}) e^{ik_T r}$ in (2.11) are defined as influence factors which describe the wave characteristics of decay and oscillation for scattering. The dominant term of an influence factor at infinity is r^{-1} and thus, the scattered field of (2.10) at infinity can be written as

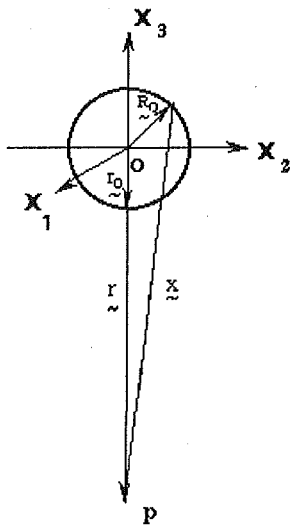


Fig. 2 Observation point *P* is at a distance from the cavity

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

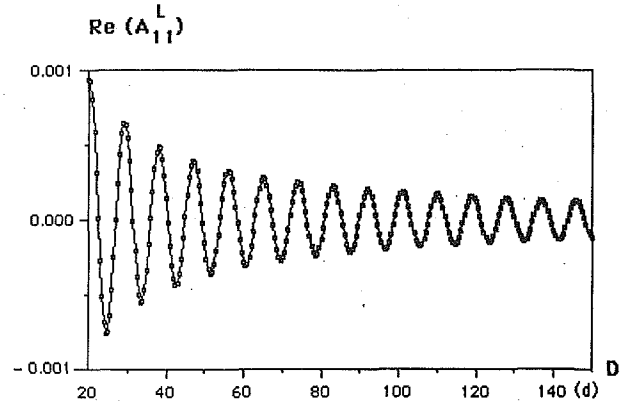


Fig. 3(*a*) The scattering influence coefficient is a function of distance *D*

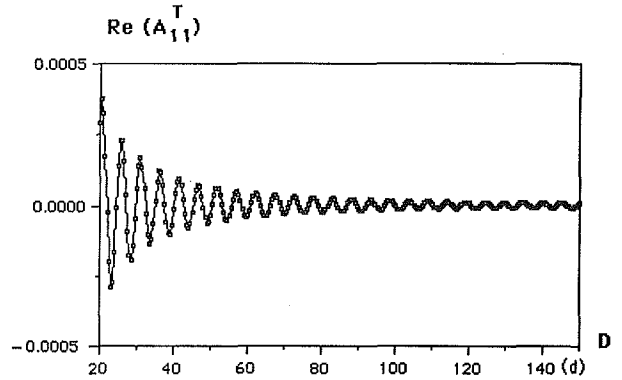


Fig. 3(*b*) The scattering influence coefficient decreases as distance increases

$$u_i^s(r) = -\frac{1}{4\pi\mu r} i\left\{\lambda\delta_{kj}\left(\frac{K_L}{K_T}\right)^2 r_{,i} K_L + \frac{2\mu}{K_T^2}(r_{,i}r_{,j}r_{,k}K_L^3)\right\} \times e^{ik_L r} V_{jk}^{L} - \frac{1}{4\pi\mu r} i\left\{\mu(\delta_{ij}r_{,k} + \delta_{ik}r_{,j})K_T - \frac{2\mu}{K_T^2}(r_{,i}r_{,j}r_{,k}K_T^3)\right\} e^{ik_T r} V_{jk}^{T}. \quad (2.13)$$

For the special case of the normal incident wave, $\mathbf{u}^I = \mathbf{p} \times \exp(iK_L\mathbf{p}\cdot\mathbf{x})$ with the vector $\mathbf{p} = (0, 0, 1)$, the back scattered field on the x_3-axis for an axially symmetric cavity (for example, a spherical cavity), can be reduced to

$$u_3^s(r) = \frac{1}{4\pi(\lambda + 2\mu)r}(iK_L) \times [\lambda(V_{11}^{L} + V_{22}^{L}) + (\lambda + 2\mu)V_{33}^{L}]e^{ik_L r} \quad (2.14)$$

where the components of the displacement u_1 and u_2 vanish because of the axial symmetry. Since the characteristics of the real part and the imaginary part of the influence function are similar, only the real parts of the influence functions are plotted in the following figures.

Figures 3(*a*) and 3(*b*) show the decay of the real components of functions $A_{11}^{L}(\mathbf{r})e^{ik_L r}$ and $A_{11}^{T}(\mathbf{r})e^{ik_T r}$ ($j = k = 1$; $i = 3$ is omitted) plotted versus D, the distance between the observation point and the center of the cavity for a normal incident wave, where the wave number is taken as $K_L d = 0.3$. Similarly, Figs. 4(*a*) and 4(*b*) ($K_L d = 0.3$) show the real components of functions $A_{33}^{L}(\mathbf{r})e^{ik_L r}$ and $A_{33}^{T}(\mathbf{r})e^{ik_T r}$ plotted versus D. It is apparent that the magnitude of the influence coefficient for the

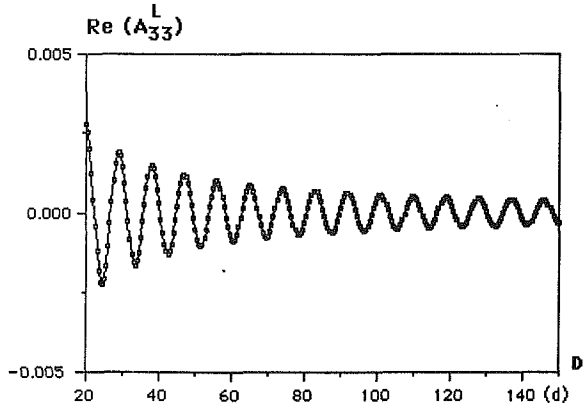


Fig. 4(*a*) The scattering influence coefficient is a function of distance *D*

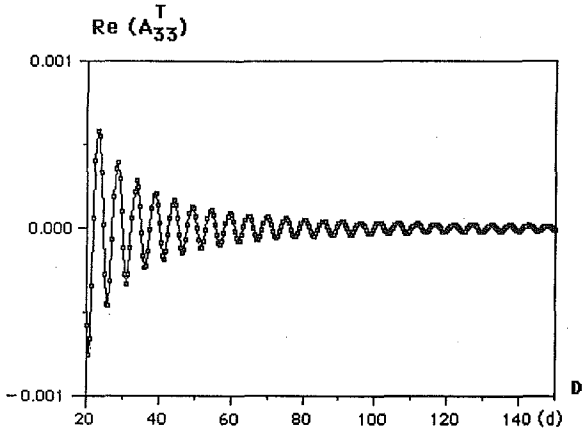


Fig. 4(*b*) The scattering influence coefficient decreases as distance increases

longitudinal wave is greater than that for the transverse wave in the x_3 direction, which implies that the cavity deformation in the direction of the incident wave (x_3) may affect the back scattered wave more than the deformation of the cavity in the direction perpendicular to x_3 does. For a group of cavities, the scattered field can be evaluated by a superposition of the COTs using Eqs. (2.11) and (2.12). The interactions of these cavity-opening tensors can be evaluated in terms of the analysis in Section 4.

3 Cavity-Opening Tensor Method for the Analysis of Reflection and Transmission

The bond plane of two solids of the same homogeneous, isotropic, and linear elastic materials is assumed to contain a large number of cavities or voids. As shown in Fig. 5, the plane

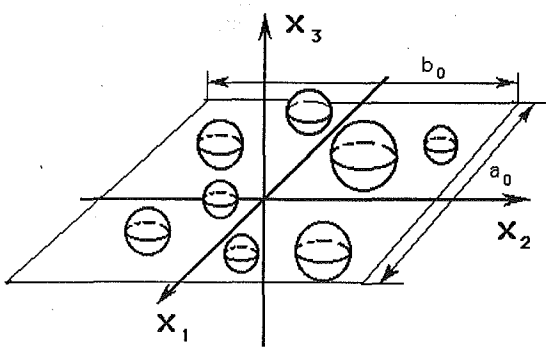


Fig. 5 Distribution of cavities in plane $x_3 = 0$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

is defined by $x_3 = 0$, and the area ($|x_1| \leq (a_0/2)$, $|x_2| \leq (b_0/2)$) in this plane intersects M cavities, which are numbered $m = 1, 2 \ldots, M$. A plane longitudinal time-harmonic wave is incident on the distribution of cavities, where $\mathbf{u}^I$ is given by

$$\mathbf{u}^I = \mathbf{p} \exp(iK_L \mathbf{p} \cdot \mathbf{x}). \tag{3.1}$$

During the bonding process, the material goes through a series of microstructural changes under certain statistically uniform manufacturing conditions. Thus, the dominant wave fields far from the bond plane may be written as reflected and transmitted plane longitudinal and transverse waves:

$x_3 \leq 0$:

$$\mathbf{u}^L = \mathbf{p} \exp(iK_L \mathbf{p} \cdot \mathbf{x}) + R^L \mathbf{p}_1 \exp(iK_L \mathbf{p}_1 \cdot \mathbf{x})$$
$$\mathbf{u}^T = R^T \mathbf{d}_1 \exp(iK_T \mathbf{p}_2 \cdot \mathbf{x})$$
$$\mathbf{u}^H = R^H \mathbf{D}_1 \exp(iK_T \mathbf{p}_2 \cdot \mathbf{x}) \tag{3.2a}$$

$x_3 \geq 0$:

$$\mathbf{u}^L = T^L \mathbf{p}_3 \exp(iK_L \mathbf{p}_3 \cdot \mathbf{x})$$
$$\mathbf{u}^T = T^T \mathbf{d}_2 \exp(iK_T \mathbf{p}_4 \cdot \mathbf{x})$$
$$\mathbf{u}^H = R^H \mathbf{D}_2 \exp(iK_T \mathbf{p}_4 \cdot \mathbf{x}) \tag{3.2b}$$

where R^L, R^T, and R^H represent the reflection coefficients for longitudinal, vertically polarized transverse and horizontally polarized transverse waves, respectively. The analogous transmission coefficients are T^L, T^T, and T^H. The vectors in (3.2a, b) are defined as

$$\mathbf{p} = \mathbf{p}_3 = (0, \sin\theta_L, \cos\theta_L), \quad \mathbf{p}_1 = (0, \sin\theta_L, -\cos\theta_L)$$
$$\mathbf{p}_2 = (0, \sin\theta_T, -\cos\theta_T), \quad \mathbf{p}_4 = (0, \sin\theta_T, \cos\theta_T)$$
$$\mathbf{d}_1 = (0, \cos\theta_T, \sin\theta_T), \quad \mathbf{d}_2 = (0, -\cos\theta_T, \sin\theta_T)$$
$$\mathbf{D}_1 = \mathbf{D}_2 = (0, 0, 1) \tag{3.3}$$

where the angles θ_L and θ_T are related by Snell's law:

$$K_L \sin\theta_L = K_T \sin\theta_T. \tag{3.4}$$

The formulation in this section is also applicable for the case of transversely incident waves as long as (3.2a) is replaced by the corresponding form of longitudinal and transverse waves.

The expressions of reflection coefficients R^α and transmission coefficients T^α, ($\alpha = L, T, H$) have been obtained by application of the reciprocal identity by Achenbach and Kitahara (1986), and the same coefficients for interface scattering were given by Xu (1992), where six auxiliary wave states which are opposite in phase and propagate in the directions opposite to the reflected and transmitted wave are applied to a cell in conjunction with the use of reciprocity. These six auxiliary waves can be expressed as $\mathbf{u}_\beta^{B\alpha}$ ($\alpha = L, T, H$; $\beta = R, T$), where α denotes three different kinds of waves, and β indicates reflection or transmission waves. Thus, the coefficients of reflection and transmission can be expressed as

$$R^\alpha = -\frac{\sum_{m=1}^{M} \int_{s_m} \mathbf{u}^A \cdot \mathbf{t}_R^{B\alpha} ds}{2iK_T \mu a_0 b_0 \cos\theta_\alpha},$$
$$T^L = \delta_{L\alpha} - \frac{\sum_{m=1}^{M} \int_{s_m} \mathbf{u}^A \cdot \mathbf{t}_T^{B\alpha} ds}{2iK_L(\lambda + 2\mu) a_0 b_0 \cos\theta_\alpha}$$
$$(\alpha = L, T, H) \tag{3.5a, b}$$

where $\mathbf{u}^A$ represents the real displacements on the cavity surfaces, and $\mathbf{t}_\beta^{B\alpha}$ ($\alpha = L, T, H$; $\beta = R, T$) is the traction corresponding to the displacement $\mathbf{u}_\beta^{B\alpha}$ for these six auxiliary waves.

It is to be noted that the reflection and transmission coefficient expressions of (3.5a, b) are based on the assumption that the average density of the COT on the diffusion plane is constant. If the distribution of the COT on this plane is completely random, the reflected and transmitted waves may have more complicated forms, and reflection and transmission coefficients would not be expressed in the explicit forms as (3.5a, b).

When the index α is taken as H, the definition $\theta_H = \theta_T$ is required. For the calculation of the reflection coefficient, the auxiliary wave $\mathbf{u}_R^{BL}$ needs to be

$$\mathbf{u}_R^{BL} = -\mathbf{p}_1 \exp(-iK_L \mathbf{p}_1 \cdot \mathbf{x}) \tag{3.6}$$

and application of Hook's law gives the traction on the cavity surface as

$$\mathbf{t}_R^{BL} = iK_L[\lambda \mathbf{n} + 2\mu(\mathbf{p}_1 \cdot \mathbf{n})\mathbf{p}_1] \exp(-iK_L \mathbf{p}_1 \cdot \mathbf{x}). \tag{3.7}$$

Subsequently, substituting (3.7) into (3.5a, b) yields the reflection coefficient

$$R^L = -\frac{1}{2a_0 b_0 \cos\theta_L} \frac{\lambda\delta_{ij} + 2\mu p_{1i} p_{1j}}{\lambda + 2\mu} \times \sum_{m=1}^{M} \int_{s_m} u_i n_j \exp(-iK_L \mathbf{p}_1 \cdot \mathbf{x}) ds \tag{3.8}$$

where p_{1i} and p_{1j} are the components of the vector $\mathbf{p}_1$. The term $-(\lambda\delta_{ij} + 2\mu p_{1i} p_{1j})/2(\lambda + 2\mu)$ is the influence coefficient and is denoted by k_{ij}^{RL}. Therefore, the explicit form of the reflection coefficient can be written as

$$R^L = k_{ij}^{RL} \cdot \frac{1}{a_0 b_0 \cos\theta_L} \sum_{m=1}^{M} V_{ij}^{m,RL}. \tag{3.9}$$

Here the term $(1/a_0 b_0 \cos\theta_L) \sum_{m=1}^{M} V_{ij}^{m,RL}$ represents the density of the COT in the plane perpendicular to the wave direction. Analogous analyses hold for the coefficients R^T, R^H, T^L, T^T, and T^H.

$$R^\alpha = k_{ij}^{R^\alpha} \cdot \frac{1}{a_0 b_0 \cos\theta_\alpha} \sum_{m=1}^{M} V_{ij}^{m,R\alpha},$$
$$T^\alpha = \delta_{L\alpha} + k_{ij}^{T^\alpha} \cdot \frac{1}{a_0 b_0 \cos\theta_\alpha} \sum_{m=1}^{M} V_{ij}^{m,T\alpha}$$
$$(\alpha = L, T, H) \tag{3.10a, b}$$

where the cavity-opening tensor of the mth cavity is

$$V_{ij}^{m,\beta\alpha} = \int_{s_m} u_i n_j \exp(-iK_\beta \mathbf{p}^{\beta\alpha} \cdot \mathbf{x}) ds,$$
$$(\alpha = L, T, H; \beta = R \cdot T) \tag{3.11}$$

which is related to the displacement over the cavity surface and to the propagation vector of the corresponding wave. The influence coefficients are

$$k_{ij}^{\beta\alpha} = -\frac{\lambda\delta_{ij} + 2\mu p_i^{\beta\alpha} p_j^{\beta\alpha}}{2(\lambda + 2\mu)}, \quad (\beta = R, T; \alpha = L) \tag{3.12}$$

$$k_{ij}^{\beta\alpha} = -\tfrac{1}{2}(d_i^{\beta\alpha} p_j^{\beta\alpha} + d_j^{\beta\alpha} p_i^{\beta\alpha}),$$
$$(\beta = R, T; \alpha = T, H). \tag{3.13}$$

The analysis above indicates that the reflection and transmission coefficients are affected by (1) distributions of the COT, which act as a series of point sources that are dominated by displacements over the cavity surfaces, and (2) influence coefficients $k_{ij}^{\beta\alpha}$, which are determined by the material properties and the direction of the incident wave.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

4 Interactions of Cavity-Opening Tensors

The value of a cavity-opening tensor of one cavity is, in general, affected by the neighboring cavities. Interactions of these cavities can be quantitatively described by using interaction factors. The following analysis is aimed at the calculation of the COTs for cavities in close proximity. The analysis will show that the interaction factors are independent, i.e., the interaction factor between two cavities depends on the locations of these two cavities, regardless of the existence of other cavities.

If closely neighboring cavities are numbered as $n = 1, 2, 3 \ldots N$, the displacements of the kth cavity can be written as

$$u_i^k = u_i^{k0} + u_i^{kN} \tag{4.1}$$

where u_i^{k0} are displacements of the kth cavity induced by the scattered field in the absence of other cavities, and u_i^{kN} are the displacement components of the kth cavity due to the presence of the other N-1 cavities. The cavity-opening tensor of cavity k, therefore, can be written as

$$V_{ij}^{k} = V_{ij}^{k0} + V_{ij}^{kN}. \tag{4.2}$$

The components of the COT due to the other cavities, V_{ij}^{kN}, can be determined using boundary integral equations as follows.

It is known that the displacement of the kth cavity, in the absence other cavities, satisfies the integral equation

$$Cu_i^{k0} = -\int_{s_k} T_{ij}^* u_j^{k0} ds + u_i^I \tag{4.3}$$

where u_i^I is the incident wave, and T_{ij}^* is Green's function. The constant C is taken as $\frac{1}{2}$ for a smooth surface s_k. For the case of N interacting cavities, the displacement of the kth cavities satisfies the integral equation

$$Cu_i^k = -\int_{s_k} T_{ij}^* u_j^k ds - \sum_{n\neq k}^{N} \int_{s_n} T_{ij}^* u_j^n ds + u_i^I. \tag{4.4}$$

Subtraction of (4.3) from (4.4) and the use of Eq. (4.1) yield

$$Cu_i^{kN} = -\int_{s_k} T_{ij}^* u_i^{kN} ds - \sum_{n\neq k}^{N} \int_{s_n} T_{ij}^* u_j^n ds. \tag{4.5}$$

It can be seen from Eqs. (4.3) and (4.5) that for the kth cavity, the displacement u_i^{k0} is determined by the incident wave u_i^I, and the displacement u_i^{kN} is determined by the displacements of the other N-1 cavities. In Eq. (4.5) there are n-1 integrals on the surfaces of the neighboring cavities, which act like n-1 sources. Therefore, the displacement u^{kN} can be decomposed into n-1 parts ($u^{kN} = \sum_{n\neq k}^{N} u_i^{kn}$) in terms of (4.5), and each of them satisfies the equation

$$Cu_i^{kn} + \int_{s_k} T_{ij}^* u_j^{kn} ds = -\int_{s_n} T_{ij}^* u_j^n ds$$

$$(n = 1, 2 \ldots N; n \neq k). \tag{4.6}$$

The integral Eq. (4.6) indicates how the displacement of cavity k is influenced by the displacement of cavity n. For convenience in finding the relation of the COTs of these two cavities, Eq. (4.6) can be written as

$$C\bar{u}_i^{kn} + \int_{s_k} \bar{T}_{ij}^* \bar{u}_j^{kn} ds = -\int_{s_n} \bar{T}_{ij}^* \bar{u}_j^n ds \tag{4.7}$$

where

$$\bar{u}_i = u_i \exp(-iK_\beta \mathbf{x}), \quad \bar{T}_{ij}^* = T_{ij}^* \exp(iK_\beta(\mathbf{y} - \mathbf{x})). \tag{4.8a, b}$$

The subscript β can be L or T, which depends on what kind of COT is needed. The source term, which is the right side of (4.7), can be approximately written as

$$-\int_{s_n} T_{ij}^*(\mathbf{x}, \mathbf{y}) u_j^n(\mathbf{y}) \exp(-iK_\beta \mathbf{p}^{\beta\alpha} \cdot \mathbf{y}) ds(\mathbf{y}) = -\sigma_{ijk}(\mathbf{x}, \mathbf{y}_0)$$

$$\times \int_{s_n} u_j^n n_k \exp(-iK_\beta \mathbf{p}^{\beta\alpha} \mathbf{y}) ds(\mathbf{y}) = -\sigma_{ijk}^*(\mathbf{x}, \mathbf{y}_0) V_{jk}^{n,\beta\alpha} \tag{4.9}$$

where $\mathbf{y}_0$ is the coordinates of the nth cavity center. Thus, Eq. (4.7) becomes

$$C\bar{u}_i^{kn} + \int_{s_k} \bar{T}_{ij} \bar{u}_j^{kn} ds = -\sigma_{ijk}^*(\mathbf{x}, \mathbf{y}_0) V_{jk}^{n,\beta\alpha}. \tag{4.10}$$

This equation indicates that cavity n, acting as a source of the COT, affects the displacements of cavity k with the influence coefficients $\sigma_{ijk}(\mathbf{x}, \mathbf{y}_0)$. It is of interest to compare the integral Eq. (4.3) for the case of a single cavity with Eq. (4.10) for the case of cavity interaction, where the displacement increment is disturbed by the source of the COT rather than u^I, the incident wave. The mathematical forms of these two equations are very similar. Equation (4.10) implies that the displacement of the kth cavity due to the presence of the nth cavity is proportional to the COT value of the nth cavity. Once the displacement increment $\bar{u}_i^{kn}$ is given from integral Eq. (4.10), the COT of the kth cavity induced by the nth cavity can be expressed as

$$V_{ij}^{kn,\beta\alpha} = \beta_{ijlm}^{kn,\beta\alpha} V_{lm}^{n,\beta\alpha}. \tag{4.11}$$

Therefore, the cavity-opening tensor of the kth cavity, in terms of (4.2), can be written as

$$V_{ij}^{k,\beta\alpha} = V_{ij}^{k0,\beta\alpha} + \sum_{n\neq k}^{N} \beta_{ijlm}^{kn,\beta\alpha} V_{lm}^{n,\beta\alpha}. \tag{4.12}$$

There are 81 interaction factor coefficients $\beta_{ijlm}^{kn,\beta\alpha}$ between two cavities because nine components of COT influence each other. Equation (4.12) exposes the relationship of the COTs between $V_{ij}^{k0,\beta\alpha}$, which is induced by the incident wave in the absence of other cavities, and $V_{ij}^{k,\beta\alpha}$, which is induced by the whole system. Once $V_{ij}^{k0,\beta\alpha}$ is given by the boundary integral equation for one cavity, $9N$ linear algebraic equations represented by (4.12) can be solved simultaneously in a straightforward manner.

Equation (4.12) also provides an alternate method, the iteration method. In the first approximation $V_{lm}^{n,\beta\alpha}$ is taken as $V_{lm}^{n0,\beta\alpha}$, the value of $V_{lm}^{n,\beta\alpha}$ in the absence of all the other cavities, i.e.,

$$V_{ij}^{k1,\beta\alpha} = V_{ij}^{k0,\beta\alpha} + \sum_{n\neq k}^{N} \beta_{ijlm}^{kn,\beta\alpha} V_{lm}^{n0,\beta\alpha} \tag{4.13}$$

and the result is substituted in the right side of (4.12). This process can be repeated as many times as necessary. This iteration method is particularly useful when the array of cavities is periodic.

5 Numerical Results of the COT Method

The cavity-opening tensor method, which treats each cavity as a point source with a three-dimensional tensor, can simplify calculations for wave scattering, wave reflection, and transmission. In general, the solution is expressed as the strength of a source, which is related to the cavity-opening tensor, multiplied by an influence function.

The scattered field from a single cavity can be expressed as a superposition of the components of a longitudinal wave and a transverse wave as shown in (2.13). The displacement on the cavity surface can be evaluated using integral Eq. (2.4) by the boundary element method. For details of the calculation one can refer to the work by Achenbach and Kitahara (1986). The quantity of the COT is calculated using

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$V_{jk}^{\alpha} = \int_s u_j n_k e^{iK_\alpha y_3} ds(\mathbf{y}). \quad (5.1)$$

For a normal incident wave by a spherical cavity, the components of V_{jk}^{α} $(j \neq K) = 0$, and the diagonal components are given by

$$V_{11}^{\alpha} = V_{22}^{\alpha} = \pi \int_0^{\pi} (u_r \sin\varphi + u_\varphi \cos\varphi) \sin^2\varphi e^{iK_\alpha \cos\varphi} d\varphi$$

$$V_{33}^{\alpha} = 2\pi \int_0^{\pi} (u_r \cos\varphi - u_\varphi \sin\varphi) \sin\varphi \cos\varphi e^{iK_\alpha \cos\varphi} d\varphi$$

$$(\alpha = L, T) \quad (5.2)$$

where u_r and u_φ are the displacements of the cavity in the polar direction and in the tangential direction, respectively. The cavity-opening tensors for a spherical cavity (192 elements on the surface) by a normal incidence $u^I = \exp(iK_L x_3)$ with the wave number $K_L d = 0.3$ are

$$\{V_{jk}^{L}\} = \begin{pmatrix} 0.020205 & & \\ & 0.020205 & \\ & & -0.038150 \end{pmatrix} + i \begin{pmatrix} 0.144458 & & \\ & 0.144458 & \\ & & 3.784920 \end{pmatrix}$$

$$\{V_{jk}^{T}\} = \begin{pmatrix} 0.019930 & & \\ & 0.019930 & \\ & & -0.009935 \end{pmatrix} + i \begin{pmatrix} 0.140769 & & \\ & 0.140769 & \\ & & 4.433268 \end{pmatrix}. \quad (5.3)$$

The back scattered field u^s can be calculated using (2.13). Figure 6(a) shows the real component of the back scattered displacement versus D (where D ranges from $20d$ to $60d$), which is the distance between the observation point and the center of the spherical cavity. The imaginary component of the back scattered displacement is plotted in Fig. 6(b), where the dot points represent the solution by the COT method and the real lines represent the solution by the boundary element method. These two curves appear to be close as the distance D becomes larger. The back scattered field with the wave number $K_L d = 1.0$ is plotted in Figs. 6(c) and 6(d). These numerical results indicate that the formulae (2.11) and (2.13) using the COT method can accurately predict the scattered field at a distance from the defects.

The COT method for the calculation of reflection and transmission of an array of cavities also has good potential applications. The reflection and transmission coefficients can, in general, be evaluated by

$$R^{\alpha} = \frac{\bar{M}}{\cos\theta_\alpha} k_{ij}^{R\alpha} \cdot \frac{1}{a_0 b_0 \cos\theta_\alpha} \bar{V}_{ij}^{R\alpha} \quad (\alpha = L, T, H)$$

$$T^{\alpha} = \delta_{\alpha L} + \frac{\bar{M}}{\cos\theta_\alpha} k_{ij}^{R\alpha} \cdot \frac{1}{a_0 b_0 \cos\theta_\alpha} \bar{V}_{ij}^{T\alpha}.$$

$$(\alpha = L, T, H) \quad (5.4a, b)$$

Here $\bar{M} = M/a_0 b_0$ denotes the number of cavity intersections per unit area in the plane $x_3 = 0$, and

$$\bar{V}_{ij}^{R\alpha} = \frac{1}{M} \sum_{m=1}^{M} V_{ij}^{R\alpha} \quad (5.5)$$

are the components of the average cavity-opening tensor. For a distribution of spherical cavities by normal incidence, the reflected coefficient $|R^L|$ versus $K_L d$ are plotted in Fig. 7. Here the Poisson's ratio of the material is selected as $\nu = 0.25$. It is of interest to compare $|R^L|$ for the case of no cavity interaction with corresponding results for the case where interaction has been taken into account. The solid line in Fig. 7 represents the rigorous results by Achenbach and Kitahara (1986) for reflection by a periodic array of spherical cavities whose centers are spaced at the distance $a/d = b/d = 3$. It can be seen that cavity

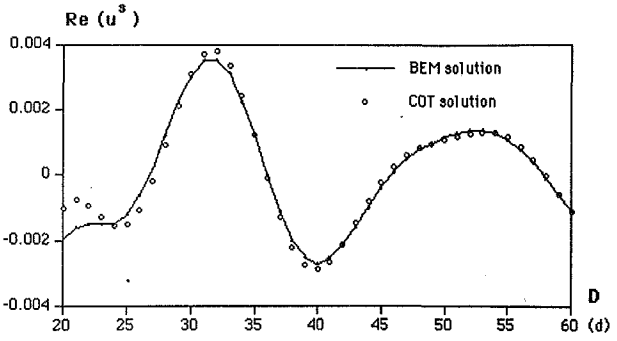


Fig. 6(a) The real part of the back scattered field ($K_L d = 0.3$)

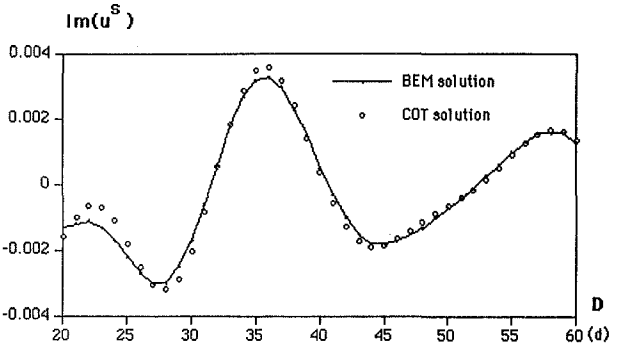


Fig. 6(b) The imaginary part of the back scattered field ($K_L d = 0.3$)

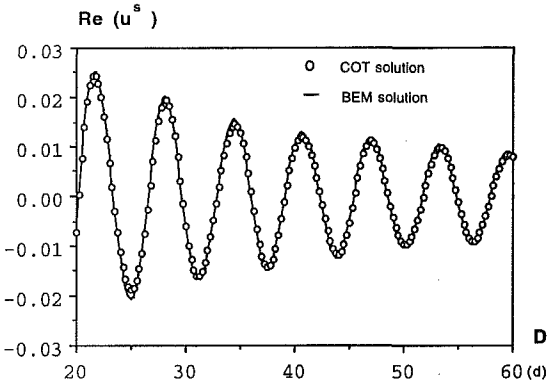


Fig. 6(c) The real part of the back scattered field ($K_L d = 1.0$)

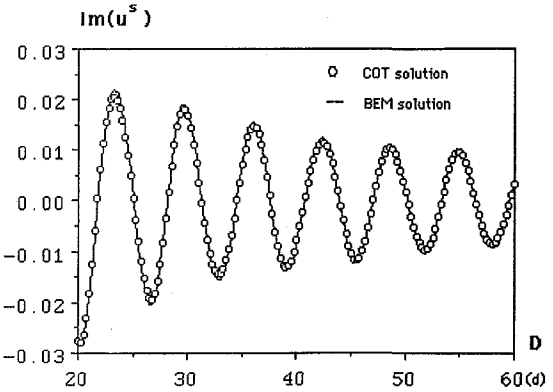


Fig. 6(d) The imaginary part of the back scattered field ($K_L d = 1.0$)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

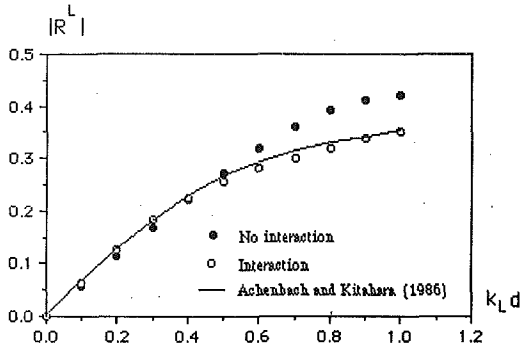


Fig. 7 Reflection coefficients for spherical cavities

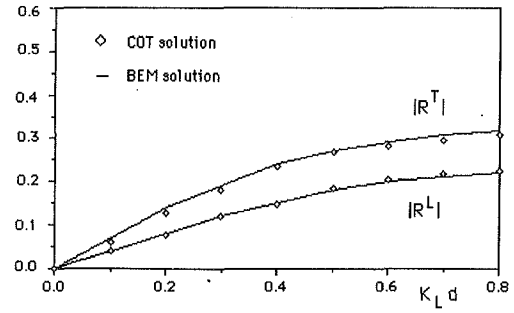


Fig. 9 Oblique incidence on periodic distribution of $a/d = 3$, $b/d = 3$

interaction starts to have a noticeable effect as the frequency becomes higher. Also plotted in Fig. 7 are the reflection and transmission coefficients for cavity interaction where the components of the COT are computed according to the interaction procedure of Eq. (4.13).

In Fig. 8 the reflection and transmission coefficients for a normal incident wave as calculated by the COT method are compared with those calculated by the boundary element method, where the cavity space distance $a/d = b/d = 4$. For the oblique incident wave the comparison of results by the COT method to those by the BEM method is given in Fig. 9. Good agreement is again observed.

6 Conclusion and Discussion

The COT method is an effective way to evaluate scattered fields and the wave reflection or wave transmission for characterizing and detecting defects in the area of nondestructive evaluation. The explicit expression of the reflection and transmission by a three-dimensional tensor makes the COT method attractive, and the interaction analysis of these tensors leads the COT method to be versatile.

The cavity opening tensor represents the deformation of the cavity. For a single cavity the COT can be calculated by the integral equation method or the boundary element method. For an array of cavities with complicated distribution, the COT method provides a way to find the interactions between cavities. The coefficients of reflected and transmitted waves can be expressed as the quantities of the COT multiplied by the influence factors. The results for calculations of the scattered wave, the reflected, and the transmitted wave are quite good as compared to the results given by rigorous numerical analysis. The COT method can be applied to the problem of cavities with more complicated distributions.

For a crack-like flaw, the tensor can be reduced to a vector for oblique incident waves or a scalar for normal incident waves. This scalar is called the crack-opening volume. The cavity-opening tenssor method can also be applied to static problems ($\omega = 0$) for the analysis of interactions between cavities and cracks.

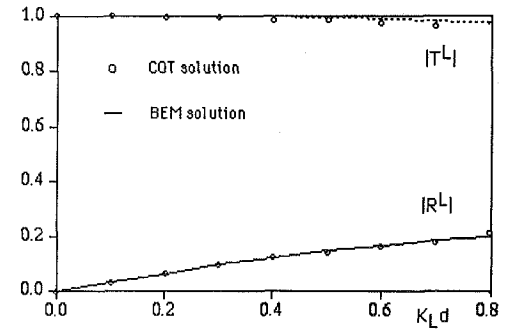


Fig. 8 Reflection and transmission coefficients for periodic distribution of $a/d = 4$, $b/d = 4$

Acknowledgment

The author is grateful to Prof. J. D. Achenbach for his helpful discussions and valuable suggestions. The support of the Office of Naval Research is gratefully acknowledged. The author would also like to thank two unidentified reviewers for their helpful comments.

References

Achenbach, J. D., and Kitahara, M., 1986, "Reflection and transmission of an obliquely incident wave by an array of a spherical cavities," *J. Acoust. Soc. Am.*, Vol. 80, pp. 1209–1214.

Achenbach, J. D., and Kitahara, M., 1987, "Hormonic waves in a solid with a periodic distribution of spherical cavities," *J. Acoust. Soc. Am.*, Vol. 81, pp. 595–598.

Achenbach, J. D., and Li, Z., 1986, "Propagation of horizontally polarized transverse waves in a solid with a periodic distribution of cracks," *Wave Motion*, Vol. 8, pp. 371–379.

Angel, Y. C., and Achenbach, J. D., 1985, "Reflection and transmission of elastic wave by a periodic array of cracks: oblique incidence," *Wave Motion*, Vol. 7, pp. 375–397.

Angel, Y. C., and Achenbach, J. D., 1987, "Hormonic waves in an elastic solid containing a doubly periodic array of cracks," *Wave Motion*, Vol. 9, pp. 377–385.

Baik, J.-M., and Thompson, R. B., 1984, "Ultrasonic scattering from imperfect interface: A quasi-static model," *J. Nondestr. Eval.*, Vol. 4, pp. 177–196.

Delph, T. J., Hermann, G., and Kaul, R. K., 1978, "Harmonic Wave Propagation in Aperiodically Layered, Infinite Elastic Body: Antiplane Strain," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 343–349.

Delph, T. J., Hermann, G., and Kaul, R. K., 1979, "Harmonic Wave Propagation in a Periodically Layered, Infinite Elastic Body: Plane Strain, Analytical Results," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 343–349.

Sotiropoulos, D. A., and Achenbach, J. D., 1988, "Reflection of elastic wave by a distribution of coplanar cracks," *J. Acoust. Soc. Am.*, Vol. 84, pp. 752–759.

Margetan, F. J., Thompson, R. B., Rose, J. H., and Gray, T. A., 1992, "The interaction of ultrasound with imperfect interface: Experimental studies of model structures," *J. Nondestr. Eval.*, Vol. 11, pp. 109–126.

Mikata, Y., 1995, "SH-Waves in a Medium Containing a Disordered Periodic Array of Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 62, pp. 312–319.

Mikata, Y., and Achenbach, J. D., 1988a, "Interaction of harmonic wave with a periodic array of inclined cracks," *Wave Motion*, Vol. 10, pp. 59–72.

Mikata, Y., and Achenbach, J. D., 1988b, "Reflection and transmission of an infinite array of randomly oriented crack," *J. Acoust. Soc. Am.*, Vol. 83, pp. 38–45.

Mikata, Y., 1993, "Reflection and Transmission by a Periodic Array of Coplanar Crack: Normal and Oblique Incidence," ASME JOURNAL OF APPLIED MECHANICS, Vol. 60, pp. 911–919.

Xu, Y., 1992, "Reflection and Transmission of Plane Wave by an Array of Cavities in the Interface of Two Solids," ASME JOURNAL OF APPLIED MECHANICS, Vol. 59, pp. 102–108.

Downloaded 04 May 2010 to 171.66.16.33. 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.

Comment on Hyperbolicity of Wave Problem for Valanis' Global Damage Theory

G. Pijaudier-Cabot[1] and Z. P. Bažant[2]

Introduction

In a very interesting and stimulating recent paper, Valanis (1991) proposed a new type of nonlocal damage theory, called the global damage theory. Some conclusions of this paper, however, require further analysis, which is the purpose of this technical note.

Offering as motivation some questionable and unfounded critical comments on the nonlocal damage theory originally proposed by Pijaudier-Cabot and Bažant (1987) and Bažant and Pijaudier-Cabot (1988), Valanis formulated a different theory and claimed it to be superior because, as he stated, it exhibits unconditional hyperbolicity of the wave problem. On closer scrutiny, though, Valanis' claim is found to be invalid, and for further development of nonlocal damage models it is important to understand why.

One-Dimensional Wave Propagation

The nature of the problem can be most simply demonstrated by considering one-dimensional wave propagation along a bar of uniform cross section. In one dimension, the constitutive relation proposed by Valanis (1991) reads:

$$\sigma = E_s\epsilon, \quad E_s = \phi^2 E_0 = (1 - D)E_0 \tag{1}$$

where σ = stress, ϵ = strain, E_0 = elastic (initial) modulus, E_s = secant modulus, D = continuum damage (see also Lemaitre, 1992), and $\phi = \sqrt{1 - D}$ = nondimensional damage parameter called by Valanis the "integrity variable"; $0 \le D \le 1$ and $0 \le \phi \le 1$. The evolution of damage is defined in a nonlocal form by

$$\dot{\phi} = -E_0\phi\epsilon^2\dot{\xi} \tag{2}$$

$$\dot{\xi}(x, t) = A\int_{V_k} k(x, x')\dot{\epsilon}(x', t)dx' \tag{3}$$

if $\dot{\epsilon} > 0$, $\epsilon > 0$, and if the integral in Eq. (3) is positive; otherwise $\dot{\xi}(x, t) = 0$. A is the cross-section area of the bar, and $k(x, x')$ is a given empirical damage influence function. Since this function must remain invariant with a translation of the coordinate system, it must have the form $k(x - x')$. This influence function is positive, continuous, and differentiable over a subset denoted as V_k of the material domain which is the integration domain for $\dot{\epsilon}(x', t)$.

Consider small perturbations of an initial equilibrium state characterized by the strain ϵ_1 and the integrity variable ϕ_1, both assumed to be initially homogeneous (uniform) over the entire bar. The equation of motion reads

$$\frac{\partial\dot{\sigma}}{\partial x} = \rho\frac{\partial^2\dot{u}}{\partial t^2} \tag{4}$$

where ρ is the mass density of the material, $\dot{\sigma}$ is the rate of stress, and $\dot{u}$ is the velocity of the small displacement perturbation. The rate of stress can be expressed as a function of the rate of strain using the constitutive relations. According to the method of linear comparison solid (Hill, 1959), we assume that the perturbation is such that $\dot{\epsilon} \ge 0$ everywhere. From Eqs. (2) and (3)

$$\dot{\phi}(x) = -E_0\phi_1\epsilon_1^2\dot{\xi}(x)$$

$$= -AE_0\phi_1\epsilon_1^2\int_{V_k} k(x - x')\frac{\partial\dot{u}(x')}{\partial x'}dx'. \tag{5}$$

From Eq. (1), the rate constitutive relation is

$$\dot{\sigma} = E_1\dot{\epsilon} + 2\dot{\phi}\phi_1E_0\epsilon_1 \tag{6}$$

where E_1 is the secant modulus at the initial state ($E_1 = \sigma_1/\epsilon_1$). Substitution of (6) into (4) requires the computation of the spatial derivative of $\dot{\phi}$. It is now important to realize that the domain V_k depends on x (note that x' is the coordinate of the material points within the integration domain). Domain V_k

[1] Professor of Civil Engineering, Laboratoire de Mécanique et Technologie, ENS Cachan/CNRS/Universite Paris VI, 61 Avenue de Président Wilson, 94235 Cachan cedex, France.

[2] Walter P. Murphy Professor of Civil Engineering, Northwestern University, Evanston, IL 60208. Mem. ASME.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME APPLIED MECHANICS DIVISION, Mar. 19, 1993; final revision, Aug. 1, 1994. Associate Technical Editor: J. G. Dvorak.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

shifts with x, therefore the spatial derivative of the integral in Eq. (5) is

$$\frac{d\left[\int_{V_k} k(x-x')\frac{\partial \dot{u}(x')}{\partial x'}dx'\right]}{dx} = \int_{V_k}\left[\frac{\partial k(x-x')}{\partial x}\frac{\partial \dot{u}(x')}{\partial x'} + \frac{\partial\left[k(x-x')\frac{\partial \dot{u}(x')}{\partial x'}\right]}{\partial x'}\right]dx'. \quad (7)$$

Alternatively, the integral term can be recast by using for the material points in V_k the coordinate $z = x' - x$ which is independent of x. Noticing that the influence function is symmetric, we thus obtain

$$\int_{V_k} k(x-x')\frac{\partial \dot{u}(x')}{\partial x}dx' = \int_{-l}^{l} k(z)\frac{\partial \dot{u}(x+z)}{\partial x}dz. \quad (8)$$

Variable z spans from $-l$ to l which is the length of domain V_k representing a basic property of the influence function. This variable is independent of the coordinate system provided the entire domain V_k is contained within the bar, and we will assume this condition in all our calculations. The calculation of the derivative of the integral in Eq. (8) with respect to x is straightforward and the result is, of course, identical to Eq. (7). Substitution of Eq. (6) with (5) and (7) into Eq. (4) yields

$$\frac{\partial^2 \dot{u}(x)}{\partial x^2} - 2AE_0\epsilon_1^2(x)\int_{-l}^{l} k(z)\frac{\partial^2 \dot{u}(x+z)}{\partial x^2}dz = \frac{1}{c_s^2}\frac{\partial^2 \dot{u}(x)}{\partial t^2} \quad (9)$$

where $c_s = \phi_1\sqrt{E_0/\rho}$ = wave velocity corresponding to the secant modulus (which represents the velocity of the unloading waves). In Valanis (1991) paper, the second term of the integrand in Eq. (7) is missing, and consequently the wave equation in Eq. (43) of that paper is not correct. Since the claim about unconditional hyperbolicity was based on this equation, that claim is also not correct. We will now examine the correct solution.

Possible Solutions

Let us now consider the harmonic wave:

$$\frac{\partial \dot{u}(x,t)}{\partial t} = Be^{-in(x-ct)} \quad (10)$$

where n is the wave number, c is the wave velocity, and $i^2 = -1$. Substituting Eq. (10) into Eq. (9), we find that the factor $Bne^{-in(x-ct)}$ cancels out from the equation and the equation of motion yields the wave velocity

$$c^2 = c_s^2[1 - 2A\epsilon_1^2\phi_1\bar{k}(n)] \quad (11)$$

with

$$\bar{k}(n) = \int_{-l}^{+l} k(z)e^{-inz}dz. \quad (12)$$

Setting $k(z) = 0$ for all $z \notin [-l, +l]$, we can regard $\bar{k}(n)$ as the Fourier transform of the damage influence function.

Equation (11) is a wave dispersion equation. This means that the velocity c is not constant but depends on the wave number n (or on the wavelength $2\pi/n$). Wave propagation in the strain-softening nonlocal material is possible as long as the expression for c^2 in Eq. (11) is real and positive. When c^2 becomes negative, wave propagation is impossible. The critical wave number n_c for which the phase velocity is zero is the solution of the equation

$$\bar{k}(n_c) = \frac{1}{2A\epsilon_1^2\phi_1}. \quad (13)$$

If the damage influence function is chosen as a bell-shaped function in the form

$$k(z) = \exp\left(\frac{|z|^2}{2l_i^2}\right) \quad (14)$$

where l_i is a characteristic length of the nonlocal continuum, the wave velocity of a small perturbation harmonic wave is

$$c^2 = c_s^2(1 - 2A\epsilon_1^2\phi_1 e^{-n^2 l_i^2/2}),$$

$$\text{with} \quad n_c = \frac{\sqrt{2}}{l_i}\sqrt{\log\left(\frac{1}{2A\epsilon_1^2\phi_1}\right)}. \quad (15)$$

Therefore, the critical wavelength $2\pi/n_c$ is proportional to the characteristic length as obtained by Pijaudier-Cabot and Bodé (1992). Furthermore, it can be checked that a critical wave number is real and greater than zero only if the material undergoes strain softening. Dispersion of waves in a nonlocal continuum is a well-known and expected property. Sluys pointed out recently (1992) the importance of dispersion in nonlocal strain-softening continua: propagation of waves with a short enough wavelength is always possible even when the material undergoes softening. This can be easily checked from Eq. (15); when $n \to \infty$ the phase velocity tends to that of the elastic unloading waves c_s, which is real.

Hyperbolicity of the Wave Equation

Ellipticity or hyperbolicity of the equation of motion cannot be decided directly. This equation is an integro-differential rather than a partial differential equation. The equation of motion can be transformed into a differential equation if the possible solution belongs to the ensemble of functions $v(x)$ for which $\partial^2 v/\partial x^2$ exists and

$$\int_{-l}^{+l} k(z)\frac{\partial^2 v(x+z)}{\partial x^2}dz = \left[\int_{-l}^{+l} k(z)F(z)dz\right]\frac{\partial^2 v(x)}{\partial x^2} \quad (16)$$

where $F(z)$ is a known function. The ensemble of such functions may not contain all the possible solutions, but the harmonic solutions considered in the previous section verify this property. For this type of functions, the equation of motion becomes

$$\frac{\partial^2 v}{\partial x^2}(x)(1 - 2AE_0\phi_1\epsilon_1^2 I^*) = \frac{1}{c_s^2}\frac{\partial^2 v}{\partial t^2},$$

$$\text{with} \quad I^* = \int_{-l}^{+l} k(z)F(z)dz. \quad (17)$$

If $v(x)$ is harmonic, I^* is the Fourier transform of the damage influence function. For this partial differential Eq. (17), the calculation of the characteristics is straightforward: the variation of the first derivatives of v is combined with the equation of motion

$$\frac{\partial^2 v}{\partial x^2}(1 - 2AE_0\phi_1\epsilon_1^2 I^*) = \frac{1}{c_s^2}\frac{\partial^2 v}{\partial t^2} \quad (18)$$

$$d\left(\frac{\partial v}{\partial t}\right) = \frac{\partial^2 v}{\partial t^2}dt + \frac{\partial^2 v}{\partial t\,\partial x}dx,$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$d\left(\frac{\partial v}{\partial x}\right) = \frac{\partial^2 v}{\partial x^2}\,dx + \frac{\partial^2 v}{\partial t\,\partial x}\,dt \tag{19}$$

and the characteristic determinant K of this system is

$$K = \frac{dx^2}{c_s^2} - (1 - 2AE_0\phi_1\epsilon_1^2 I^*)dt^2. \tag{20}$$

The characteristic lines in the (x, t) space are given by the equation

$$\left(\frac{dx}{dt}\right) = c_s\sqrt{(1 - 2AE_0\phi_1\epsilon_1^2 I^*)} \tag{21}$$

which yields, for harmonic waves, exactly the wave velocity.

Hyperbolicity of the partial differential Eq. (17) is lost when the characteristic lines lie in the imaginary region, i.e., when the wave number is bellow the critical wave number for harmonic waves defined in Eq. (13). Therefore, we conclude that the wave equation is not unconditionally hyperbolic for Valanis' nonlocal model. When $v(x)$ is harmonic (a possible admissible solution), the integro-differential equation of motion (9) reduces to an elliptic partial differential equation if wave number is bellow the critical wave number. However, nonlocality of the type considered in this note enables waves with a short wave length to propagate in the softening regime (Pijaudier-Cabot and Benallal, 1992).

Conclusions

(1) Wave propagation in a medium described by the global damage theory is dispersive. This feature is restricted neither to Valanis' theory nor to the integral-type nonlocal constitutive relations in general. As pointed out by Sluys (1992), all the constitutive models proposed in the literature that incorporate an internal length are dispersive: the micro-polar continua, the gradient-dependent models and also rate-dependent models. For these models, Sluys demonstrated that hyperbolicity of the wave equation is always preserved.

(2) The wave equation of Valanis' global damage theory is not unconditionally hyperbolic. This result is at variance with the assertion in Valanis' paper (1991) which was based on an incorrect wave equation. The derivative of the damage controlling variable $\xi(x, t)$ with respect to coordinate x was not calculated taking into account the variation of the limits of the nonlocal averaging integral. The question of well-posedness is, however, still an open issue since finite bodies and initial conditions ought to be included in such considerations.

Acknowledgment

Partial financial support of the Commission of the European Communities through the Brite-Euram Programme (Project BE-3275) to the first author is gratefully acknowledged. The work of the second author was partially supported under AFOSR grant 91-0140 to Northwestern University.

References

Bažant, Z. P., and Pijaudier-Cabot, G., 1988, "Nonlocal Continuum Damage: Localization, Instability, and Convergence," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 287–293.

Hill, R., 1959, "Some Basic Principles in the Mechanics of Solids without a Natural Time," *Journal of Mechanics and Physics of Solids,* Vol. 7, pp. 209–225.

Lemaitre, J., 1992, *A course on Damage Mechanics,* Springer-Verlag, New York.

Pijaudier-Cabot, G., and Bažant, Z. P., 1987, "Nonlocal Damage Theory," *ASCE Journal of Engineering Mechanics,* Vol. 113, pp. 1512–1533.

Pijaudier-Cabot, G., and Benallal, A., 1993, "Strain Localization and Bifurcation in a Nonlocal Continuum," *International Journal of Solids and Structures,* Vol. 30, pp. 1761–1775.

Pijaudier-Cabot, G., and Bodé, L., 1992, "Localization of Damage in a Nonlocal Continuum," *Mechanics Research Communications,* Vol. 19, pp. 145–153.

Sluys, L. J., 1992, "Wave Propagation, Localization and Dispersion in Softening Solids," Doctoral dissertation, Technical University of Delft, The Netherlands.

Valanis, K. C., 1991, "A Global Damage Theory and the Hyperbolicity of the Wave Problem," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 311–316.

Hypersingular Integral Equation for a Curved Crack Problem of Circular Region in Antiplane Elasticity

Y. Z. Chen[3] and Norio Hasebe[4]

1 Introduction

To solve the crack problem, the hypersingular integral equation approach was developed recently by Kaya and Erdogan (1987), Nied (1987), and Ioakimidis (1988). In this approach, the finite-part integral proposed by Hadamard (1923) is a foundation for analysis. However, the previous investigations are limited to the case of straight crack configuration. A curved crack problem of circular region in antiplane elasticity is considered in this note. Taking the displacement jump function along the curved crack to be unknown function and the traction to be the right-side term in the resulting integral equation, and using the image method, a hypersingular integral equation for the problem is formulated. One numerical example with the calculated results is presented.

2 Definition of the Hypersingular Integral Along a Curve

Let the configuration of the curve L be expressed by $t(s) = x(s) + iy(s)$, where s is the arc length in Fig. 1, and both functions $x(s)$ and $y(s)$ have continuous first and second derivatives with respect to s. Secondly, we assume that the function $G(t)$ is a bounded function with first and second continuous derivatives and is defined along a smooth curve L. In this case, the hypersingular integral along the curve L may be defined as

$$\int_L \frac{G(t)dt}{(t - t_0)^2} = \int_L \frac{G(t) - G(t_0) - (t - t_0)G'(t_0)}{(t - t_0)^2}\,dt + G'(t_0)\int_L \frac{dt}{t - t_0} + G(t_0)\int_L \frac{dt}{(t - t_0)^2} \quad (t_0 \in L). \tag{1}$$

The first integral on the right-hand side of (1) is an integral

[3] Division of Engineering Mechanics, Jiangsu University of Science and Technology, Zhengiang, Jaingsu, 212013, P.R. China.

[4] Department of Civil Engineering, Nagoya Institute of Technology, Gokisocho, Showa-ku, Nagoya 466, Japan.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 20, 1993; final revision, Mar. 9, 1996. Associate Technical Editor: D. M. Parks.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$d\left(\frac{\partial v}{\partial x}\right) = \frac{\partial^2 v}{\partial x^2}\,dx + \frac{\partial^2 v}{\partial t\,\partial x}\,dt \tag{19}$$

and the characteristic determinant K of this system is

$$K = \frac{dx^2}{c_s^2} - (1 - 2AE_0\phi_1\epsilon_1^2 I^*)dt^2. \tag{20}$$

The characteristic lines in the (x, t) space are given by the equation

$$\left(\frac{dx}{dt}\right) = c_s\sqrt{(1 - 2AE_0\phi_1\epsilon_1^2 I^*)} \tag{21}$$

which yields, for harmonic waves, exactly the wave velocity.

Hyperbolicity of the partial differential Eq. (17) is lost when the characteristic lines lie in the imaginary region, i.e., when the wave number is bellow the critical wave number for harmonic waves defined in Eq. (13). Therefore, we conclude that the wave equation is not unconditionally hyperbolic for Valanis' nonlocal model. When $v(x)$ is harmonic (a possible admissible solution), the integro-differential equation of motion (9) reduces to an elliptic partial differential equation if wave number is bellow the critical wave number. However, nonlocality of the type considered in this note enables waves with a short wave length to propagate in the softening regime (Pijaudier-Cabot and Benallal, 1992).

Conclusions

(1) Wave propagation in a medium described by the global damage theory is dispersive. This feature is restricted neither to Valanis' theory nor to the integral-type nonlocal constitutive relations in general. As pointed out by Sluys (1992), all the constitutive models proposed in the literature that incorporate an internal length are dispersive: the micro-polar continua, the gradient-dependent models and also rate-dependent models. For these models, Sluys demonstrated that hyperbolicity of the wave equation is always preserved.

(2) The wave equation of Valanis' global damage theory is not unconditionally hyperbolic. This result is at variance with the assertion in Valanis' paper (1991) which was based on an incorrect wave equation. The derivative of the damage controlling variable $\xi(x, t)$ with respect to coordinate x was not calculated taking into account the variation of the limits of the nonlocal averaging integral. The question of well-posedness is, however, still an open issue since finite bodies and initial conditions ought to be included in such considerations.

Acknowledgment

Partial financial support of the Commission of the European Communities through the Brite-Euram Programme (Project BE-3275) to the first author is gratefully acknowledged. The work of the second author was partially supported under AFOSR grant 91-0140 to Northwestern University.

References

Bažant, Z. P., and Pijaudier-Cabot, G., 1988, "Nonlocal Continuum Damage: Localization, Instability, and Convergence," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 287–293.

Hill, R., 1959, "Some Basic Principles in the Mechanics of Solids without a Natural Time," *Journal of Mechanics and Physics of Solids,* Vol. 7, pp. 209–225.

Lemaitre, J., 1992, *A course on Damage Mechanics,* Springer-Verlag, New York.

Pijaudier-Cabot, G., and Bažant, Z. P., 1987, "Nonlocal Damage Theory," *ASCE Journal of Engineering Mechanics,* Vol. 113, pp. 1512–1533.

Pijaudier-Cabot, G., and Benallal, A., 1993, "Strain Localization and Bifurcation in a Nonlocal Continuum," *International Journal of Solids and Structures,* Vol. 30, pp. 1761–1775.

Pijaudier-Cabot, G., and Bodé, L., 1992, "Localization of Damage in a Nonlocal Continuum," *Mechanics Research Communications,* Vol. 19, pp. 145–153.

Sluys, L. J., 1992, "Wave Propagation, Localization and Dispersion in Softening Solids," Doctoral dissertation, Technical University of Delft, The Netherlands.

Valanis, K. C., 1991, "A Global Damage Theory and the Hyperbolicity of the Wave Problem," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 311–316.

Hypersingular Integral Equation for a Curved Crack Problem of Circular Region in Antiplane Elasticity

Y. Z. Chen[3] and Norio Hasebe[4]

1 Introduction

To solve the crack problem, the hypersingular integral equation approach was developed recently by Kaya and Erdogan (1987), Nied (1987), and Ioakimidis (1988). In this approach, the finite-part integral proposed by Hadamard (1923) is a foundation for analysis. However, the previous investigations are limited to the case of straight crack configuration. A curved crack problem of circular region in antiplane elasticity is considered in this note. Taking the displacement jump function along the curved crack to be unknown function and the traction to be the right-side term in the resulting integral equation, and using the image method, a hypersingular integral equation for the problem is formulated. One numerical example with the calculated results is presented.

2 Definition of the Hypersingular Integral Along a Curve

Let the configuration of the curve L be expressed by $t(s) = x(s) + iy(s)$, where s is the arc length in Fig. 1, and both functions $x(s)$ and $y(s)$ have continuous first and second derivatives with respect to s. Secondly, we assume that the function $G(t)$ is a bounded function with first and second continuous derivatives and is defined along a smooth curve L. In this case, the hypersingular integral along the curve L may be defined as

$$\int_L \frac{G(t)dt}{(t - t_0)^2} = \int_L \frac{G(t) - G(t_0) - (t - t_0)G'(t_0)}{(t - t_0)^2}\,dt + G'(t_0)\int_L \frac{dt}{t - t_0} + G(t_0)\int_L \frac{dt}{(t - t_0)^2} \quad (t_0 \in L). \tag{1}$$

The first integral on the right-hand side of (1) is an integral

[3] Division of Engineering Mechanics, Jiangsu University of Science and Technology, Zhengiang, Jaingsu, 212013, P.R. China.

[4] Department of Civil Engineering, Nagoya Institute of Technology, Gokisocho, Showa-ku, Nagoya 466, Japan.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 20, 1993; final revision, Mar. 9, 1996. Associate Technical Editor: D. M. Parks.

Copyright © 1996 by ASME

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

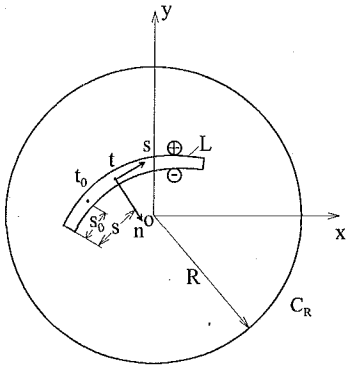


Fig. 1 A curved crack in the circular region

in the usual sense. The second integral can be understood as a Cauchy principal value integral (Muskhelishvili, 1954).

Alternatively, the hypersingular integral can be defined as

$$\int_L \frac{G(t)dt}{(t-t_0)^2} = \frac{d}{dt_0}\int_L \frac{G(t)dt}{t-t_0} \quad (t_0 \in L). \tag{2}$$

Using this definition, the third integral on the right-hand side of (1) can be evaluated:

$$\int_L \frac{dt}{(t-t_0)^2} = \frac{d}{dt_0}\int_L \frac{dt}{t-t_0} = \frac{1}{c-t_0} - \frac{1}{d-t_0} \quad (t_0 \in L). \tag{3}$$

3 Formulation of Hypersingular Integral Equation for a Curved Crack in a Circular Region

It is well known that after using the complex potential $\phi(z)$ in antiplane elasticity, we have (Chen and Wang, 1986)

$$\phi(z) = Gw(x, y) + if(x, y) \tag{4}$$

$$f(x, y) = \int^{(x,y)} (\sigma_{xz}dy - \sigma_{yz}dx) \tag{5}$$

$$\phi'(z) = \sigma_{xz} - i\sigma_{yz} = G\left(\frac{\partial w}{\partial x} - i\frac{\partial w}{\partial y}\right) = \frac{\partial f}{\partial y} + i\frac{\partial f}{\partial x} \tag{6}$$

$$\sigma_{nz} = G\frac{\partial w}{\partial n} = \frac{\partial f}{\partial s} \tag{7}$$

where G is the shear modulus of elasticity, w is the longitudinal displacement, f is the longitudinal resultant force, and σ_{xz}, σ_{yz}, and σ_{nz} are the stress components.

Consider the cracked round bar in antiplane shear loading (Fig. 1). Without loss of generality, we assume that the circular boundary C_R is traction-free and the curved crack L is subjected to some longitudinally distributed traction $p(s)$. Thus, the governing equation and the boundary conditions for the function $w(x, y)$ can be expressed as

$$\nabla^2 w(x, y) = 0, \left(\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \tag{8}$$

$$\left.\frac{\partial w}{\partial n}\right|_{C_R} = 0, \quad \text{(along the outer circular boundary } C_R) \tag{9}$$

$$\left.\frac{\partial f}{\partial s}\right|_L = G\left.\frac{\partial w}{\partial n}\right|_L = p(s) \quad \text{(along the curved crack } L). \tag{10}$$

The appropriate complex potential for the curved crack in an infinite region takes the form

$$\phi(z) = \frac{i}{\pi}\int_L \frac{D(t)dt}{t-z} \tag{11}$$

where $D(t)$ is a real density function depending on t ($t \in L$).

Noting the property $z\bar{z} = R^2 (z \in C_R)$ and using the image method (Chen, 1991), the complex potential for the curved crack of the circular region will be

$$\phi(z) = Gw(x, y) + if(x, y) = \frac{i}{\pi}\left[\int_L \frac{D(t)dt}{t-\bar{z}} + \int_L \frac{zD(t)\overline{dt}}{R^2 - \bar{t}z}\right]. \tag{12}$$

It is easy to prove that the condition (9) is satisfied by the function $w(x, y)$ in (12).

After using the Plemelj formula (Muskhelishvili, 1954), we have

$$\phi^+(t_0) - \phi^-(t_0) = -2D(t_0) \quad (t_0 \in L). \tag{13}$$

This is to say, the function $w(x, y)$ is discontinuous and the function $f(x, y)$ is continuous when the point z is across the curve L.

Letting $z \to t_0^+$ or $z \to t_0^-$ and taking the imaginary part of (12), the boundary condition (10) can be rewritten as

$$\mathrm{Re}\left[\int_L \frac{D(t)dt}{t-t_0} + \int_L \frac{t_0 D(t)\overline{dt}}{R^2 - \bar{t}t_0}\right] = \pi\int_0^{s_0} p(s)ds + c \quad (t_0 \in L). \tag{14}$$

Differentiating both sides of (14) with respect to s_0, the following hypersingular integral equation is obtainable:

$$\mathrm{Re}\left[\frac{dt_0}{ds_0}\int_L \frac{D(t)dt}{(t-t_0)^2} + \frac{dt_0}{ds_0}\int_L \frac{R^2}{(R^2 - \bar{t}t_0)^2}D(t)\overline{dt}\right] = \pi p(s_0). \tag{15}$$

If a round torsion bar is weakened by a curved crack L (Fig. 1), the governing equation and the one boundary condition for the warping function $w(x, y)$ can be also referred to (8) and (9), and the boundary condition along the curved crack takes the form (Sokolnikoff, 1956)

$$\left.\frac{\partial w}{\partial n}\right|_L = y\cos(x, n) - x\cos(y, n) \quad \text{(along the curved crack } L, (x, y) \in L). \tag{16}$$

Thus, the formulation proposed in the curved crack problem of antiplane elasticity is still valid in the present case.

4 Numerical Example

In the numerical solution, the curved crack configuration is approximated by the winding crack consisting of N line segments as shown in Fig. 2. Two points are important to model

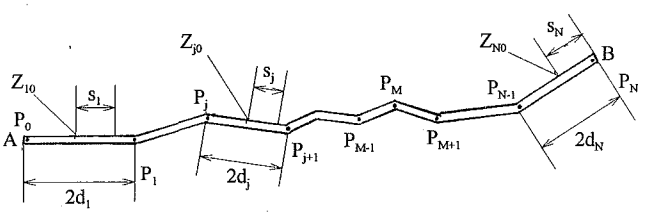


Fig. 2 A winding crack consisting of N line segments

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the COD (crack opening displacement) function in the solution. First, the adopted COD function should meet the physical situation of the deformation geometry, particularly at the vicinity of crack tip. Secondly, the assumed COD function will make the computation possible. It is found that the appropriate COD function $D(s)$ can be assumed in the following manner:

(1) For the first segment in Fig. 2, we choose

$$D(s_1) = c_1 g(s_1), \quad g(s_1) = \sqrt{\frac{1}{2}\left(1 + \frac{s_1}{d_1}\right)} \quad |s_1| \le d_1 \quad (17)$$

The function $g(s_1)$ can reflect the nature of deformation at the vicinity of crack tip. Since the function $g(s_1)$ is not convenient for integration, and it is approximated by

$$g(s_1) \approx h(s_1) = \left(1 - \frac{s_1^2}{d_1^2}\right)^{1/2}\left(1 - \frac{s_1}{d_1}\right)\left(\alpha - \beta\frac{s_1}{d_1}\right) + \frac{1}{2}\left(1 + \frac{s_1}{d_1}\right) \quad |s_1| \le d_1, \quad (18)$$

the parameters α and β involved are determined by

$$\lim_{s_1 \to -d_1} g(s_1) \Big/ \left(1 + \frac{s_1}{d_1}\right)^{1/2} = \lim_{s_1 \to d_1} h(s_1) \Big/ \left(1 + \frac{s_1}{d_1}\right)^{1/2} \quad (19)$$

and

$$g(s_1)|_{s_1 = d_1/5} = h(s_1)|_{s_1 = d_1/5}. \quad (20)$$

The condition (19) means the two functions have same deformation nature at the crack tip. In addition the condition (20) means the two functions take the same value at the weight coordinate $s_1 = d_1/5$. From (19) and (20), it follows that

$$\alpha = \frac{1}{4} - \beta \quad \beta = \frac{5}{24} - \frac{125}{48\sqrt{6}}\left(\sqrt{\frac{3}{5}} - \frac{3}{5}\right). \quad (21)$$

(2) For the intermediate segments in Fig. 2, we choose

$$D(s_j) = \frac{c_j}{2}\left(1 - \frac{s_j}{d_j}\right) + \frac{c_{j+1}}{2}\left(1 + \frac{s_j}{d_j}\right) \quad |s_j| < d_j \quad (j = 2, 3, \ldots, N-1). \quad (22)$$

(3) Similar to the first segment, for the last segment we can let

$$D(s_N) = c_{N-1}\left\{\left(1 - \frac{s_N^2}{d_N^2}\right)^{1/2}\left(1 + \frac{s_N}{d_N}\right)\left(\alpha + \beta\frac{s_N}{d_N}\right) + \frac{1}{2}\left(1 - \frac{s_N}{d_N}\right)\right\} |s_N| < d_N. \quad (23)$$

In this case, the number of total unknowns is $N - 1$. Meantime, the integral equation is reduced to

$$\mathrm{Re}\left[\sum_{j=1}^{N}\left\{\frac{dt_0}{ds_0}\int_{L_j}\frac{D(t)dt}{(t - t_0)^2} + \frac{dt_0}{ds_0}\int_{L_j}\frac{R^2}{(R^2 - \bar{t}t_0)^2}D(t)\,\overline{dt}\right\}\right] = \pi p(s_0) \quad (24)$$

where L_j means the integration performed along the jth segment. In computation, we choose the middle points of the segment to be the collocation points, and these points are denoted by Z_{10}, Z_{20}, ... Z_{N0} in Fig. 2. Therefore, N algebraic equations are available after discretation. To balance the numbers of the unknowns and the equations, we reduce two equations into one equation simply by adding two equations together.

We can see later the integrals involved in Eq. (24) can be reduced into two types of integral, and be integrated in a closed form. For example, we consider the first integral in (24),

$$I = \frac{dt_0}{ds_0}\int_{L_j}\frac{D(t)dt}{(t - t_0)^2}. \quad (25)$$

There are two particular cases for our consideration:

(1) In the case of $t_0 \in L_j$ (Fig. 3(a)), we can make the following substitution in (25):

$$t \Rightarrow Z_{j0} + s_j e^{i\alpha_j}, \; t_0 \Rightarrow Z_{j0} + s_{j0}e^{i\alpha_j}, \quad dt \Rightarrow e^{i\alpha_j}ds_j, \; \frac{dt_0}{ds_0} \Rightarrow e^{i\alpha_j} \quad (26)$$

and obtain

$$I_1 = \int_{-d_j}^{d_j}\frac{D(s_j)ds_j}{(s_j - s_{j0})^2} \quad |s_{j0}| < d_j. \quad (27)$$

This is precisely a hypersingular integral.

(2) In the case of $t_0 \notin L_j$ (Fig. 3(b)), we can make the following substitution in (25):

$$t \Rightarrow Z_{j0} + s_j e^{i\alpha_j}, \quad dt \Rightarrow e^{i\alpha_j}ds_j, \quad \frac{dt_0}{ds_0} \Rightarrow e^{i\beta} \quad (28)$$

and obtain

$$I_2 = e^{i(\beta - \alpha_j)}\int_{-d_j}^{d_j}\frac{D(s_j)ds_j}{(s_j - p)^2}, \quad p = e^{-i\alpha_j}(t_0 - Z_{j0}). \quad (29)$$

Clearly, this is a regular integral.

Similar analysis can be made to the second integral in Eq. (24), However, in this case we reduce the investigated integrals to the regular one as shown by (29).

Finally, we can reduce all the investigated integrals in the form

$$J_1(s_0) = \int_{-d}^{d}\frac{D(s)ds}{(s - s_0)^2} \quad |s_0| < d \quad (30)$$

$$J_2(p) = \int_{-d}^{d}\frac{D(s)ds}{(s - p)^2} \quad p \notin (-d, d). \quad (31)$$

The positions of s_0 and the point p are indicated in Fig. 3(c).

From (17), (18), (22), and (23) we see that the used COD functions are the linear combination of the following functions:

$$D(s) = \left(1 - \frac{s^2}{d^2}\right)^{1/2}, \quad D(s) = \left(1 - \frac{s^2}{d^2}\right)^{1/2}\left(\frac{s}{d}\right), \quad D(s) = \left(1 - \frac{s^2}{d^2}\right)^{1/2}\left(\frac{s}{d}\right)^2,$$

$$D(s) = 1, \quad D(s) \doteq \frac{s}{d}. \quad (32)$$

Substituting the functions shown by Eq. (32) into (30) and (31) and using the known results (Chen, 1993) yields

$$\int_{-d}^{d}\left(1 - \frac{s^2}{d^2}\right)^{1/2}\frac{ds}{(s - s_0)^2} = -\frac{\pi}{d}$$

$$\int_{-d}^{d}\left(1 - \frac{s^2}{d^2}\right)^{1/2}\left(\frac{s}{d}\right)\frac{ds}{(s - s_0)^2} = -\frac{\pi}{d^2}(2s_0)$$

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

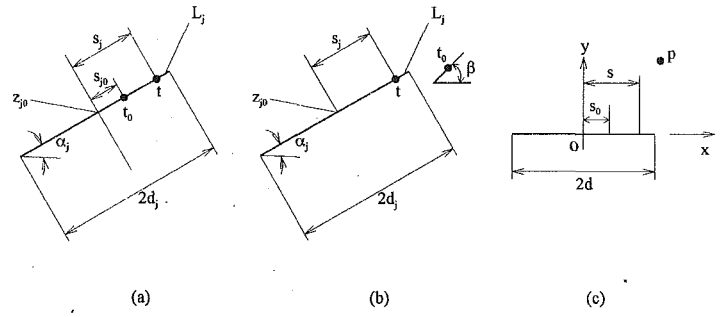


Fig. 3 (*a*) a segment L_j for evaluating the integral shown by (27), $t_0 \in L_j$; (*b*) a segment L_j for evaluating the integral shown by (29), $t_0 \notin L_j$; (*c*) a segment $(-d, d)$ for evaluating the hypersingular and regular integrals

$$\int_{-d}^{d} \left(1 - \frac{s^2}{d^2}\right)^{1/2} \left(\frac{s}{d}\right)^2 \frac{ds}{(s - s_0)^2} = -\frac{\pi}{d^3}\left(3s_0^2 - \frac{d^2}{2}\right)$$

$$\int_{-d}^{d} \frac{ds}{(s - s_0)^2} = -\frac{2d}{(d^2 - s_0^2)}$$

$$\int_{-d}^{d} \frac{s}{d} \frac{ds}{(s - s_0)^2} = \frac{1}{d}\left\{\log(d - s_0) - \log(d + s_0) - \frac{2ds_0}{d^2 - s_0^2}\right\} \quad (33)$$

$$\int_{-d}^{d} \left(1 - \frac{s^2}{d^2}\right)^{1/2} \frac{ds}{(s - p)^2} = -\frac{\pi}{d}\left(1 - \frac{p}{X(p)}\right)$$

$$\int_{-d}^{d} \left(1 - \frac{s^2}{d^2}\right)^{1/2} \left(\frac{s}{d}\right) \frac{ds}{(s - p)^2} = -\frac{\pi}{d^2}\left[2(p - X(p)) - \frac{d^2}{X(p)}\right]$$

$$\int_{-d}^{d} \left(1 - \frac{s^2}{d^2}\right)^{1/2} \left(\frac{s}{d}\right)^2 \frac{ds}{(s - p)^2} = -\frac{\pi}{d^3}\left[3p(p - X(p)) - \frac{d^2}{2} - \frac{pd^2}{X(p)}\right]$$

$$\int_{-d}^{d} \frac{ds}{(s - p)^2} = -\frac{2d}{(d^2 - p^2)}$$

$$\int_{-d}^{d} \left(\frac{s}{d}\right) \frac{ds}{(s - p)^2} = \frac{1}{d}\left[\log(d - p) - \log(-d - p) - \frac{2dp}{d^2 - p^2}\right] \quad (34)$$

where

$$X(p) = \sqrt{p^2 - d^2}$$

(taking the branch $\lim_{z \to \infty} X(z)/z = 1$). (35)

Finally, we can conclude that after using the suggested interpolation formulae for the COD function, all the integrals involved can be quadratured in a closed form. Furthermore, after the algebraic equation for $c_1, c_2, \ldots c_{N-1}$ is solved, we can get the stress intensity factor by

$$K_{3A} = \sqrt{2\pi} \lim_{s_1 \to -d_1} D'(s_1)\sqrt{d_1 + s_1}$$

(for the left crack tip in Fig. 2)

$$K_{3B} = \sqrt{2\pi} \lim_{s_N \to d_N} D'(s_N)\sqrt{d_N - s_N}$$

(for the right crack tip in Fig. 2). (36)

A numerical example is carried out in this study. In the example, a circular arc crack is embedded in the circular region with radius R, and the uniform longitudinal traction q is applied on the crack face (Fig. 4). We use $N = 20$ in computation (N-

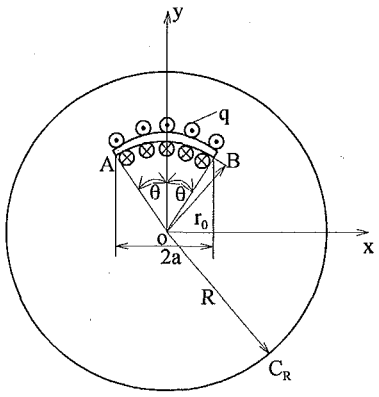


Fig. 4 A circular arc crack embedded in circular region

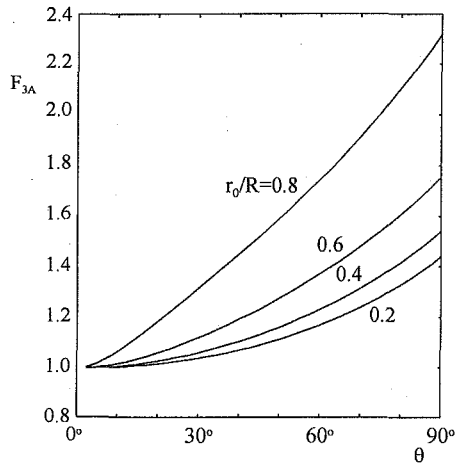


Fig. 5 The calculated stress intensity factors (see (37) and Fig. 4)

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

number of the divisions), and express the calculated stress intensity factor by

$$K_{3A} = F_{3A}\left(\theta, \frac{r_0}{R}\right) q\sqrt{\pi a} \quad (a = r_0 \sin\theta). \tag{37}$$

The calculated F_{3A} values are plotted in Fig. 5. From Fig. 5 we see that if the circular arc crack approaches the outer boundary, the stress intensity factors become larger.

References

Chen, Y. Z., and Wang, Z. X., 1986, "Solutions of Multiple Crack Problems of a Circular Region with Free or Fixed Boundary Condition in Antiplane Elasticity," *International Journal of Fracture,* Vol. 30, pp. 287–293.

Chen, Y. Z., 1991, "Image Method for Curved Crack Problem in Antiplane Elasticity," *International Journal of Fracture,* Vol. 48, pp. R75–R78.

Chen, Y. Z., 1993, "Numerical Solution of a Curved Crack Problem by Using Hypersingular Integral Equation Approach," *Engineering Fracture Mechanics,* Vol. 46, pp. 275–283.

Hadamard, J., 1923, *Lectures on Cauchy's Problem in Linear Partial Differential Equations,* Yale University Press, New Haven, CT.

Ioakimidis, N. I., 1988, Mangler-type Principal Value Integrals in Hypersingular Integral Equations for Crack Problems in Plane Elasticity," *Engineering Fracture Mechanics,* Vol. 31, pp. 895–989.

Kaya, A. C., and Erdogan, F., 1987, "On the Solution of Integral Equations with Strongly Singular Kernels," *Quart. Appl. Math.,* Vol. 45, pp. 105–122.

Muskhelishvili, N. I., 1954, *Some Basic Problems of Mathematical Theory of Elasticity,* Noordhoff, Netherlands.

Nied, H. F., 1987, "Periodic Array of Cracks in a Half-plane Subjected to Arbitrary Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 542–648.

Sokolnikoff, S., 1956, *Mathematical Theory of Elasticity,* McGraw-Hill, New York.

Downloaded 04 May 2010 to 171.66.16.33. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm