Transactions of the ASME®

Technical Editor, LEON M. KEER (1993)

The Technological Institute
Northwestern University
Evanston, IL 60201

APPLIED MECHANICS DIVISION

Chairman, MICHAEL J. FORRESTAL
Secretary, L. B. FREUND
Associate Technical Editors,
R. M. BOWEN (1994)
S. K. DATTA (1993)
G. J. DVORAK (1993)
M. E. FOURNEY (1994)
C. O. HORGAN (1992)
R. L. HUSTON (1994)
D. J. INMAN (1992)
J. T. C. LIU (1993)
R. M. McMEEKING (1994)
A. K. NOOR (1994)
J. W. RUDNICKI (1992)
W. N. SHARPE, JR. (1994)
C. F. SHIH (1993)
J. G. SIMMONDS (1992)
P. D. SPANOS (1992)
K. R. SREENIVASAN (1992)
Z. WARHAFT (1992)

BOARD ON COMMUNICATIONS
Chairman and Vice-President
R. E. NICKELL
Members-at-Large
W. BEGELL, T. F. CONRY, T. DEAR,
R. L. KASTOR, R. MATES,
E. M. PATTON, R. E. REDER,
R. D. ROCKE, A. VAN DER SLUYS,
A. J. WENNERSTROM, W. O. WINER,
B. ZIELS

President, A. E. BERGLES
Exec. Dir.
D. L. BELDEN
Treasurer, ROBERT A. BENNETT

PUBLISHING STAFF
Mng. Dir., Publ.,
CHARLES W. BEARDSLEY
Managing Editor,
CORNELIA MONAHAN
Production Editor, JUDY SIERANT
Prod. Asst., MARISOL ANDINO

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $160.00 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. 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, $29.00, annually; to nonmembers, $160.00. Add $20.00 for postage to countries outside the United States and Canada.

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

COPYRIGHT © **1991** by The American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the TRANSACTIONS OF THE ASME, JOURNAL OF APPLIED MECHANICS, and the author, and date of publication be stated.

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc.

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 58 • NUMBER 1 • MARCH 1991

TECHNICAL PAPERS

1 Dynamic Spherical Cavity Expansion of Strain-Hardening Materials
V. K. Luk, M. J. Forrestal, and D. E. Amos

7 Penetration of Strain-Hardening Targets With Rigid Spherical-Nose Rods
M. J. Forrestal, N. S. Brar, and V. K. Luk

11 Acoustoelastic Birefringences in Plastically Deformed Solids: Part I—Theory (90-WA/APM-41)
Yih-Hsing Pao, Tsung-Tsong Wu, and Udo Gamer

18 Acoustoelastic Birefringences in Plastically Deformed Solids: Part II—Experiment (90-WA/APM-42)
Tsung-Tsong Wu, Masahiko Hirao, and Yih-Hsing Pao

24 Calculations of Strains and Internal Displacement Fields Using Computerized Tomography (91-APM-14)
R. J. Gu, J. D. Hovanesian, and Y. Y. Hung

28 A New Boundary Element Method for Three-Dimensional Coupled Problems of Consolidation and Thermoelasticity
G. F. Dargush and P. K. Banerjee

37 Fatigue, Fractals, and a Modified Miner's Rule
Eli Altus

43 Statistical Theory for Predicting the Failure of Brittle Materials (91-WA/APM-1)
S. She, J. D. Landes, J. A. M. Boulet, and J. E. Stoneking

50 The Proportional Anisotropic Elastic Invariants (91-APM-13)
A. M. Sadegh and S. C. Cowin

58 Stress-Focusing Effect in a Uniformly Heated Solid Sphere
Toshiaki Hata

64 Thermal Shock in a Hollow Sphere Caused by Rapid Uniform Heating
Toshiaki Hata

70 Creasing Singularities in Compressible Elastic Materials
S. A. Silling

75 An Edge Dislocation in a Three-Phase Composite Cylinder Model
H. A. Luo and Y. Chen

87 On the Extent of Dominance of Asymptotic Elastodynamic Crack-Tip Fields: Part I—An Experimental Study Using Bifocal Caustics
Sridhar Krishnaswamy and Ares J. Rosakis

95 On the Extent of Dominance of Asymptotic Elastodynamic Crack-Tip Fields: Part II—Numerical Investigation of Three-Dimensional and Transient Effects
Sridhar Krishnaswamy, Ares J. Rosakis, and G. Ravichandran

104 Two-Parameter Characterization of Elastic-Plastic Crack-Tip Fields
C. Betagón and J. W. Hancock

111 Near-Tip Fields for Quasi-Static Crack Growth Along a Ductile-Brittle Interface
W. J. Drugan

120 The Hertzian Cone Crack
Li Yingzhi and D. A. Hills

128 Rigid Punch on Stressed Orthotropic Half-Plane With Partial Slip (91-APM-11))
H. L. Bjarnehed

134 Complementarity Problem Formulation of Three-Dimensional Frictional Contact
B. M. Kwak

141 A Thermal Elastica Theory of Piston-Ring and Cylinder-Bore Contact (91-APM-8)
D. C. Sun

154 Pure Bending of Curved Beams of Thin-Walled Rectangular Box Section
R. D. Cook

157 Local Delamination Buckling in Layered Systems (90-WA/APM-40)
E. Madenci and R. A. Westmann

167 Buckling Calculations and Measurements on a Technologically Relevant Toroidal Magnet System
W. Geiger and K. P. Jüngst

175 The Usefulness of Elementary Theory for the Linear Vibrations of Layered, Orthotropic Elastic Beams and Corrections Due to Two-Dimensional End Effects
J. M. Duva and J. G. Simmonds

181 Equations of Elastic Deformation of Laminated Composite Shallow Shells (91-APM-25)
A. W. Leissa and M. S. Qatu

189 Active Vibration Control of the Axially Moving String in the *S* Domain (91-WA/APM-2)
B. Yang and C. D. Mote, Jr.

197 Optimal Control of Beams With Dynamic Loading and Buckling (91-APM-5)
Y.-A. Su and I. G. Tadjbakhsh

203 Accurate Analytical-Type Solutions for the Free Vibration of Simply-Supported Parallelogram Plates (91-APM-6)
D. J. Gorman

(Contents continued on Inside Back Cover)

Downloaded 04 May 2010 to 171.66.16.49. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

CONTENTS (CONTINUED)

209 **Bending Frequency of a Rotating Beam With an Elastically Restrained Root** (91-APM-7)
Sen Yung Lee and Yee Hsiung Kuo

215 **Dynamics of Multibody Systems With Known Configuration Changes**
J. J. McPhee and R. N. Dubey

222 **Free Vibration of Helical Springs** (91-APM-4)
W. Jiang, W. K. Jones, T. L. Wang, and K. H. Wu

229 **Stability of Linear Conservative Gyroscopic Systems**
J. A. Walker

233 **An Inverse Problem Associated With Modification of Incomplete Dynamic Systems**
Y. M. Ram and S. G. Braun

238 **Dynamics of Nonholonoic Mechanical Systems Using a Natural Orthogonal Complement**
Subir Kumar Saha and Jorge Angeles

244 **Chaos in a Weakly Nonlinear Oscillator With Parametric and External Resonances** (91-APM-3)
K. Yagasaki

251 **A Modified, Slowly Varying Parameter Approach for Systems With Impulsive Loadings**
Nicholas P. Jones and Harry W. Shenton III

259 **Co-dimensional Two Bifurcations in the Presence of Noise**
N. Sri Namachchivaya

266 **Maximum Entropy Approach for Stationary Response of Nonlinear Stochastic Oscillators**
R. J. Chang

272 **Inertia Effects in Thin Film Flow With a Corrogated Boundary**
Ilter Serbetci and John A. Tichy

278 **Stability of Liquid Film Flow Down an Oscillating Wall**
Ronald J. Bauer and C. H. von Kerzcek

BRIEF NOTES

Universal Relations for Elastically Isotropic Elastic-Plastic Materials — 283
M. B. Rubin and R. Chen

Response of a Crack to Transient Concentrated Line Forces — 285
Chen Baoxing and Zhang Xiangzhou

The Shear Modulus of Foamlike Structures — 288
D. Stamenović

On the Shear Modulus of Polyhedron-Cell Liquid Foam — 289
B. Budiansky and E. Kimmel

290 On the Crack Simulation and Solutions in the Lattice
N. F. Morozov and M. W. Paucshto

292 Bending of Plates With Rigid Inclusions Having Cuspidal Points
E. E. Gdoutos, M. A. Kattis, and M. Katsarou

294 On the Competition Between Failure and Instability in Progressively Softening Solids
A. Needleman

296 The Motion of a Sphere on a Rough Horizontal Plane
J.-Y. Yang

BOOK REVIEWS

299 ***Computational Mechanics of Probabilistic and Reliability Analysis*, by W. K. Liu and T. B. Belytschko** . . . Reviewed by **P. Spanos**

299 ***Non-Destructive Testing*, by B. Hull and V. John** . . . Reviewed by **Toshio Mura**

ANNOUNCEMENTS AND SPECIAL NOTES

36 Complas III—Call for Papers

49 First U.S. National Congress on Computational Mechanics—Announcement

57 ASTM International Symposium on Automation in Fatigue and Fracture Testing and Analysis—Call for Papers

86 Change of Address Form

166 28th Meeting of the Society of Engineering Science—Call for Papers

228 18th International Congress of Theoretical and Applied Mechanics—Call for Papers

300 I.U.T.A.M. Symposium—Call for Proposals

301 Books Received by the Office of the Technical Editor

302 Worldwide Mechanics Meetings List

304 Information for Authors

Downloaded 04 May 2010 to 171.66.16.49. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

V. K. Luk
Mem. ASME

M. J. Forrestal
Fellow ASME

D. E. Amos

Sandia National Laboratories,
Albuquerque, NM 87185

Dynamic Spherical Cavity Expansion of Strain-Hardening Materials[1]

We developed models for the dynamic expansion of spherical cavities from zero initial radii for elastic-plastic, rate-independent materials with power-law strain hardening. The models considered the material as incompressible and compressible. For an incompressible material, we obtained closed-form solutions, whereas the compressible results required the numerical solution of differential equations. A comparison of the numerical results from both models showed the effect of compressibility.

Introduction

Hopkins (1960) extensively reviewed the expansion of spherical cavities for elastic-plastic materials. Briefly, Bishop, Hill, and Mott (1945) derived equations for the quasi-static expansion of cylindrical and spherical cavities for metals, and Chadwick (1959) investigated ideal soils. Hill (1948) also derived equations for the dynamic expansion of an incompressible, elastic-plastic material. Later, Hunter and Crozier (1968) and Forrestal and Luk (1988) studied the dynamic cavity expansion of a compressible, elastic, perfectly plastic solid. While Hopkins (1960) discussed several practical applications that motivated the solution of cavity expansion problems, our interest is in penetration mechanics (Forrestal, Okajima, and Luk, 1988). Hill (1980) also applied cavity expansion methods to investigate armor penetration by nondeforming projectiles.

Previous models for the dynamic expansion of spherical cavities took the material as incompressible with linear strain, hardening (Hill, 1948 and 1950) or as a compressible, elastic perfectly plastic material (Forrestal and Luk, 1988). Kawahara (1986) recently developed a compression test for ductile metals to obtain stress-strain data to 100 percent true strain at rates to 10^2 s^{-1}. As will be shown, a power-law data-fit closely approximated the post-yield data obtained by Kawahara. Motivated by these data, we developed models for the dynamic expansion of spherical cavities from zero initial radii for elastic-plastic materials with power-law strain hardening. The models considered the material as incompressible and compressible. For an incompressible material, we obtained closed-form solutions, whereas the compressible results required the numerical solution of differential equations. A comparison of the numerical results from both models showed the effect of compressibility.

Problem Formulation

A spherically symmetric cavity is expanded from zero initial radius at constant velocity V. As shown in Fig. 1, this expansion produces plastic and elastic response regions. The plastic region is bounded by the radii $r = Vt$ and $r = ct$, where r is the radial, Eulerian coordinate, t is time, and c is the elastic-plastic interface velocity. For the compressible model the elastic region is bounded by $r = ct$ and $r = c_d t$, where c_d is the elastic, dilatational velocity. For the incompressible model, there is no elastic wave front, and the elastic region occupies $r > ct$.

Plastic Response. Equations of momentum and mass conservation in Eulerian coordinates with spherical symmetry are

$$\frac{\partial \sigma_r}{\partial r} + \frac{2(\sigma_r - \sigma_\theta)}{r} = -\rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial r}\right) \tag{1a}$$

$$\rho_0 \frac{\partial}{\partial r}[(r - u)^3] = 3\rho r^2 \tag{1b}$$

where σ_r, σ_θ are the radial and hoop components of Cauchy

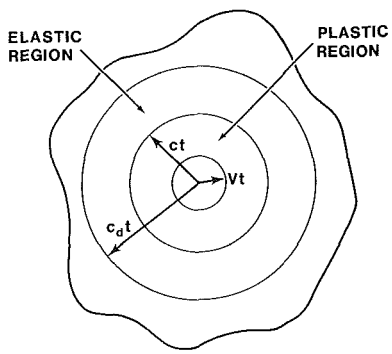


Fig. 1 Response regions

[1]Presented at the 11th U.S. National Congress of Applied Mechanics, Tucson, Ariz., May 21–25, 1990.

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

Discussion of this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 3, 1989; final revision, Jan. 19, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

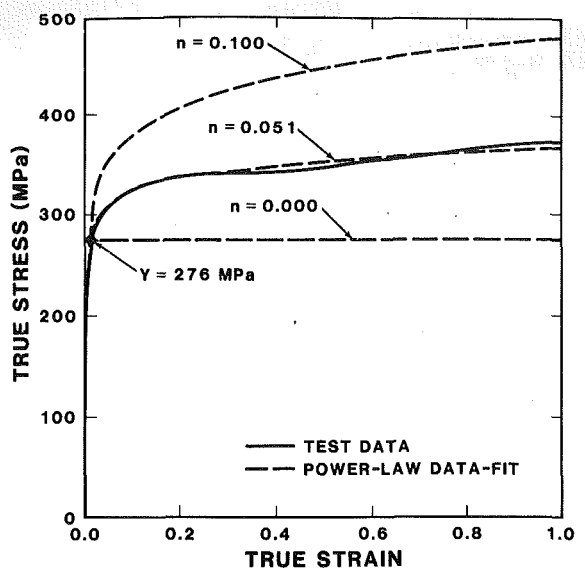


Fig. 2 True stress versus true strain for 6061-T651 aluminum and the power-law data-fits with E = 68.9 GPa, ν = 1/3, and Y = 276 MPa

stress, measured positive in compression; ρ_0 and ρ are the densities in the undeformed and deformed states; and u and v are the displacement and particle velocity in the radial direction, with outward motion considered positive. Particle velocity and displacement are related by

$$v\left(1 - \frac{\partial u}{\partial r}\right) = \frac{\partial u}{\partial t}. \tag{2}$$

In this paper, we develop solutions for incompressible and compressible materials. For an incompressible material, $\rho = \rho_0$ in equations (1*a*, *b*). Description of the compressible material, however, requires the pressure-volumetric strain relation

$$p = 1/3\ (\sigma_r + 2\sigma_\theta) = K\ (\epsilon_r + 2\epsilon_\theta) \tag{3a}$$

$$K = E/[3(1 - 2\nu)] \tag{3b}$$

where K is the bulk modulus, ϵ_r and ϵ_θ are logarithmic strain components, E is Young's modulus, and ν is Poisson's ratio. Plastic flow is introduced by the method presented by Hill (1950) for the quasi-static problem. For spherical symmetry, the principal stresses (σ_r, σ_θ, σ_θ) consist of the isotropic state (σ_θ, σ_θ, σ_θ) and the state of uniaxial stress in the radial direction ($\sigma_r - \sigma_\theta$, 0, 0). Thus, we can apply the data from a uniaxial compression test, provided that we allow for the elastic strains produced by the isotropic stress state. We use the modified Ludwik equation (Chakrabarty, 1987) as the power-law data-fit to describe the post-yield, uniaxial stress-strain data. Thus,

$$\sigma = E\epsilon,\ \sigma \le Y \tag{4a}$$

$$\sigma = Y(E\epsilon/Y)^n,\ \sigma > Y \tag{4b}$$

where σ is Cauchy stress (true stress), ϵ is logarithmic strain (true strain), Y is the yield stress, and n is the strain-hardening exponent. Figure 2 shows the power-law data-fit to 6061-T651 aluminum data.

The uniaxial radial stress ($\sigma_r - \sigma_\theta$, 0, 0) consists of the principal stress state from cavity expansion (σ_r, σ_θ, σ_θ) that produces radial strain ϵ_r and the isotropic stress ($-\sigma_\theta$, $-\sigma_\theta$, $-\sigma_\theta$) that produces radial strain $[-(1 - 2\nu)\sigma_\theta/E]$. Therefore, the uniaxial radial stress ($\sigma_r - \sigma_\theta$, 0, 0) produces the total strain $[\epsilon_r - (1 - 2\nu)\sigma_\theta/E]$. With the aid of equation (4*b*), plastic flow for the cavity expansion problem is governed by

$$\sigma_r - \sigma_\theta = E^n\ Y^{1-n}\ [\epsilon_r - (1 - 2\nu)\sigma_\theta/E]^n. \tag{5}$$

The logarithmic strain-displacement relations are

$$\epsilon_r = \ln\ (1 - \partial u/\partial r) \tag{6a}$$

$$\epsilon_\theta = \ln\ (1 - u/r). \tag{6b}$$

Boundary and Elastic-Plastic Interface Conditions. The boundary condition at the cavity surface is

$$u(r = Vt) = Vt. \tag{7}$$

The plastic and elastic regions of response are linked through the Hugoniot jump conditions that conserve mass and momentum at the elastic-plastic interface. The Tresca yield criterion,[2] $\sigma_r - \sigma_\theta = Y$, applies on either side of the interface, and Forrestal and Luk (1988) proved that radial stress and particle velocity are continuous at the interface.

Elastic Response. Forrestal and Luk (1988) presented elastic response equations for both incompressible and compressible models.

Incompressible Strain-Hardening Material

As pointed out by Hill (1950), the analysis can be simplified by assuming that the material is incompressible. For an incompressible material $\rho = \rho_0$ and $\nu = 1/2$. Thus, density is known and the second strain term in equation (5) vanishes.

Plastic Region. Using equations (5) and (6*a*), the momentum and mass conservation equations (1*a*), (1*b*) reduce to

$$\frac{\partial \sigma_r}{\partial r} + \frac{2}{r} E^n\ Y^{1-n}\left[\ln\left(1 - \frac{\partial u}{\partial r}\right)\right]^n = -\rho_0\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial r}\right) \tag{8}$$

$$\frac{\partial}{\partial r}[(r - u)^3] = 3r^2. \tag{9}$$

With equation (2) that relates u and v, we have three equations and three unknowns: σ_r, u, and v. We introduce the similarity transformation

$$\xi = r/ct \tag{10a}$$

and the dimensionless variables

$$S = \sigma_r/Y,\ \bar{u} = u/ct,\ U = v/c,\ \gamma = V/c. \tag{10b}$$

With equations (10*a*), (10*b*), equations (2), (8), and (9) transform to

$$\frac{dS}{d\xi} = -\frac{2}{\xi}\left(\frac{E}{Y}\right)^n\left[\ln\left(1 - \frac{d\bar{u}}{d\xi}\right)\right]^n + \frac{\rho_0 c^2}{Y}(\xi - U)\frac{dU}{d\xi} \tag{11a}$$

$$\frac{d}{d\xi}[(\xi - \bar{u})^3] = 3\xi^2 \tag{11b}$$

$$U\left(1 - \frac{d\bar{u}}{d\xi}\right) = \left(\bar{u} - \xi\frac{d\bar{u}}{d\xi}\right). \tag{11c}$$

The boundary condition equation (7) transforms to

$$\bar{u}\ (\xi = \gamma) = \gamma. \tag{12}$$

A solution of equation (11*b*) that satisfies equation (12) is

$$\bar{u} = \xi - (\xi^3 - \gamma^3)^{1/3},\ \gamma \le \xi \le 1. \tag{13}$$

From equations (13) and equations (11*c*), the dimensionless particle velocity is

$$U = \gamma^3/\xi^2,\ \gamma \le \xi \le 1 \tag{14}$$

where $\xi = 1$ is the elastic-plastic interface location. Substitution of equations (13) and (14) into (11*a*), and integrating gives

[2]Tresca and Mises yield criteria are equivalent for spherical symmetry (Mendelson, 1968).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$S = S_2 + \frac{2}{3}\left(\frac{2E}{3Y}\right)^n I(\xi) - \frac{\rho_0 c^2 \gamma^3}{2Y}\left(4 - \frac{4}{\xi} + \frac{\gamma^3}{\xi^4}\right), \quad \gamma \le \xi \le 1 \tag{15a}$$

where $S_2 = S(\xi = 1)$ is the dimensionless radial stress at the interface, and

$$I(\xi) = \int_{1-(\gamma/\xi)^3}^{1-\gamma^3} \frac{(-\ln x)^n}{1-x}\, dx. \tag{15b}$$

The integral $I(\xi)$ in equation (15*b*) is obtained through the successive transformation $\zeta = (\gamma/\xi)^3$ and $x = (1 - \zeta)$. When $n = 1$, $I(\xi)$ is the difference between two dilogarithm functions (Abramowitz and Stegun, 1964). Amos (1988) presents a numerical procedure to evaluate $I(\xi)$.

Interface Conditions. Forrestal and Luk (1988) show that stress and particle velocity are continuous at the interface, so

$$S_2 = S_1,\ U_2 = U_1 \tag{16}$$

where subscripts 2 and 1 refer to quantities at the interface in the plastic and elastic regions, respectively.

Elastic Region. The solutions for the elastic region (Forrestal and Luk, 1988) are

$$\bar{u} = \frac{Y}{2E} \cdot \frac{1}{\xi^2},\ \xi \ge 1 \tag{17a}$$

$$U = \frac{3Y}{2E} \cdot \frac{1}{\xi^2},\ \xi \ge 1 \tag{17b}$$

$$S = \frac{2}{3} \cdot \frac{1}{\xi^3} + \frac{3\rho_0 c^2}{E} \cdot \frac{1}{\xi},\ \xi \ge 1. \tag{17c}$$

Response Equations. The interface velocity c is obtained by equating particle velocities at the interface equation (16). Thus,

$$c = \left(\frac{2E}{3Y}\right)^{1/3} V. \tag{18}$$

With equations (16), (17), and (18), the radial stress in the plastic region becomes

$$S(\xi) = \frac{2}{3} + \frac{2}{3}\left(\frac{2E}{3Y}\right)^n I(\xi) + \frac{\rho_0 V^2}{Y}\left(\frac{3Y}{2E}\right)^{1/3}\left(\frac{2}{\xi} - \frac{3Y}{4E} \cdot \frac{1}{\xi^4}\right), \quad \gamma \le \xi \le 1. \tag{19}$$

The radial stress at the cavity surface $\xi = \gamma$ reduces to

$$S(\xi = \gamma) = A + BV^2 \tag{20a}$$

where

$$A = \frac{2}{3}\left[1 + \left(\frac{2E}{3Y}\right)^n I(\xi = \gamma)\right] \tag{20b}$$

$$B = \frac{3}{2}\frac{\rho_0}{Y} \tag{20c}$$

The dynamic term BV^2 is the same as that derived by Forrestal and Luk (1988) for an elastic, perfectly plastic material. Therefore, strain hardening affects only the quasi-static solution given by A in equation (20*b*).

Compressible Strain-Hardening Material

For a compressible material, the density ρ is a dependent variable, and the total strain in equation (5) must be kept in the analysis. The compressibility equation that relates volumetric strain to pressure in equation (3*a*) is also required for the compressible solution.

Plastic Region. The plastic flow rule equation (5), with the aid of equations (3*a*), (3*b*) and the strain definitions equations (6*a*), (6*b*), is written as

$$\sigma_r - \sigma_\theta = \left(\frac{E}{2Y}\right)^n Y\left\{\ln\left[\left(1 - \frac{\partial u}{\partial r}\right)\left(\frac{r}{r-u}\right)^2\right] + \frac{\sigma_r}{3K}\right\}^n. \tag{21}$$

The compressibility equation (3*a*), when combined with the plastic flow rule equation (21) to eliminate σ_θ, can be expressed as

$$\sigma_r - \frac{2}{3}Y\left(\frac{E}{2Y}\right)^n\left\{\ln\left[\left(1 - \frac{\partial u}{\partial r}\right)\left(\frac{r}{r-u}\right)^2\right] + \frac{\sigma_r}{3K}\right\}^n - K\ln\left[\left(1 - \frac{\partial u}{\partial r}\right)\left(1 - \frac{u}{r}\right)^2\right] = 0. \tag{22}$$

For the compressible material, we normalize radial stress by K and keep the other definitions given by equations (10*a*), (10*b*). We define the additional, dimensionless variables

$$S = \sigma_r/K,\ T = Y/K,\ \lambda = c/c_p,\ \phi = E/2Y \tag{23}$$

where $c_p^2 = K/\rho_0$. With equations (2), (10*a*), (10*b*), (21), and (23), the momentum and mass conservation equations (1*a*), (1*b*) transform to

$$\frac{dS}{d\xi} + \frac{2T\phi^n G^n\left(\xi, \bar{u}, \frac{d\bar{u}}{d\xi}, S\right)}{\xi} = \left(\frac{\rho}{\rho_0}\right)\lambda^2(\xi - U)\frac{dU}{d\xi} \tag{24a}$$

$$(\xi - \bar{u})^2\left(1 - \frac{d\bar{u}}{d\xi}\right) = \left(\frac{\rho}{\rho_0}\right)\xi^2 \tag{24b}$$

where

$$G\left(\xi, \bar{u}, \frac{d\bar{u}}{d\xi}, S\right) = \ln\left[\left(1 - \frac{d\bar{u}}{d\xi}\right)\left(\frac{\xi}{\xi - \bar{u}}\right)^2\right] + \frac{S}{3}. \tag{24c}$$

Equations (24*a*), (24*b*) are combined with equation (11*c*) to eliminate U and ρ. The resulting equation and equation (22) provide two equations for the unknown, S and $\bar{u}$. Thus,

$$\frac{dS}{d\xi} + \frac{2T\phi^n G^n\left(\xi, \bar{u}, \frac{d\bar{u}}{d\xi}, S\right)}{\xi} = -\frac{\lambda^2(\xi - \bar{u})^4}{\xi^2\left(1 - \frac{d\bar{u}}{d\xi}\right)^2} \cdot \frac{d^2\bar{u}}{d\xi^2} \tag{25a}$$

$$S - \frac{2}{3}T\phi^n G^n\left(\xi, \bar{u}, \frac{d\bar{u}}{d\xi}, S\right) - \ln\left[\left(1 - \frac{d\bar{u}}{d\xi}\right)\left(1 - \frac{\bar{u}}{\xi}\right)^2\right] = 0. \tag{25b}$$

Equations (25*a*), (25*b*) are further simplified by introducing a new variable,

$$\omega = \xi - \bar{u} \tag{26}$$

that converts the Eulerian coordinates to the Lagrangian system. Equations (25*a*), (25*b*) become

$$\frac{dS}{d\xi} + \frac{2T\phi^n G^n\left(\xi, \omega, \frac{d\omega}{d\xi}, S\right)}{\xi} = \left(\frac{\lambda\omega^2}{\xi\frac{d\omega}{d\xi}}\right)^2 \cdot \frac{d^2\omega}{d\xi^2} \tag{27a}$$

$$S - \frac{2}{3}T\phi^n G^n\left(\xi, \omega, \frac{d\omega}{d\xi}, S\right) - \ln\left[\frac{d\omega}{d\xi}\left(\frac{\omega}{\xi}\right)^2\right] = 0 \tag{27b}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

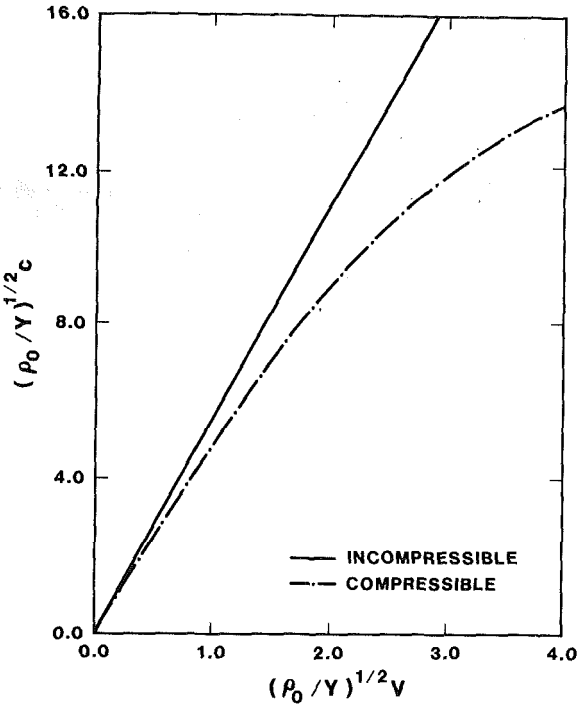


Fig. 3 Elastic-plastic interface velocity versus cavity expansion velocity for $n = 0.051$

where

$$G(\xi, \omega, \frac{d\omega}{d\xi}, S) = \ln\left[\left(\frac{\xi}{\omega}\right)^2 \frac{d\omega}{d\xi}\right] + \frac{S}{3}. \tag{27c}$$

The three conditions at the interface required to solve the coupled differential equations (27*a*), (27*b*) are

$$S(\xi = 1) = S_2 \tag{28a}$$

$$\omega(\xi = 1) = (1 - U_2)\frac{d\omega}{d\xi}(\xi = 1) \tag{28b}$$

$$S_2 - \frac{2}{3} T \phi^n G^n(\xi = 1) - \ln\left[\frac{d\omega}{d\xi}(\xi = 1)\,\omega^2(\xi = 1)\right] = 0 \tag{28c}$$

where S_2 and U_2 are the dimensionless radial stress and particle velocity in the plasic region at the interface $\xi = 1$, equation (28*b*) is obtained from $U(\xi = 1) = U_2$ with equations (11*c*) and (26), while equation (28*c*) is obtained from equation (27*b*). The boundary condition at the cavity surface is found from equations (12) and (26) and given by

$$\omega(\xi = \gamma) = 0. \tag{29}$$

We now have four conditions to solve numerically for the four unknowns ω, $\frac{d\omega}{d\xi}$, S, and λ in equations (27*a*), (27*b*).

Interface Conditions. At the elastic-plastic interface, $\sigma_r - \sigma_\theta = Y$, and the material has not strain hardened. Therefore, radial stress and particle velocity are continuous at the interface $\xi = 1$ (Forrestal and Luk, 1988). Thus, equation (16) applies.

Elastic Region. Forrestal and Luk (1988) present solutions for the elastic region. Only solutions at the elastic-plastic interface ($\xi = 1$) are summarized. Thus,

$$S_1 = \frac{2T}{3}\left[1 + \frac{(1 + \nu)\zeta^2}{(1 - 2\nu)(1 + \zeta)}\right], \xi = 1 \tag{30a}$$

$$\bar{u}_1 = \frac{T(1 + \nu)}{9(1 - 2\nu)} \cdot \left(1 - \frac{2\zeta^2}{1 + \zeta}\right), \xi = 1 \tag{30b}$$

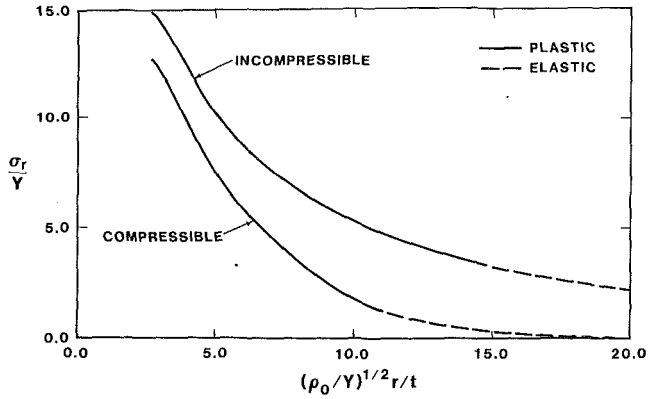


Fig. 4 Radial stress profiles for $(\rho_0/Y^{1/2})$ $V = 2.70$ and $n = 0.051$

$$U_1 = \frac{T(1 + \nu)}{3(1 - 2\nu)}, \xi = 1 \tag{30c}$$

where

$$\zeta = c/c_d, c_d^2 = \frac{3(1 - \nu)}{(1 + \nu)} c_p^2. \tag{30d}$$

Response Equations. The boundary conditions in equations (28*a*), (28*b*), and (28*c*), when linked with the elastic solutions in equations (30*a*), (30*c*) through the continuity interface conditions equation (16), provide the necessary conditions to solve numerically the differential equations in equations (27*a*), (27*b*). The numerical solution procedures are described in the next subsection.

Numerical Solution Procedure. The thrust of the numerical procedure is to solve equations (27*a*), (27*b*) not only for the distributions S and ω, but also for γ using equation (29). Thus, we select a value of λ, $0 < \lambda < 1$, and start a backward integration of equations (27*a*), (27*b*) into the region $0 < \gamma \leq \xi \leq 1$ with initial conditions at $\xi = 1$ formed from equations (28*b*), (28*c*), and (30*a*). When $\omega = 0$, γ is known, and the expansion velocity $V = \gamma c$ can be computed.

The set of equations in equations (27*a*), (27*b*) is a differential algebraic system. While there are numerical integration routines designed for this type of system (Petzold, 1982; Stepelman, 1983), it is more convenient for most users to differentiate equation (27*b*) and use a conventional integrator that solves the system in a canonical form. With this method, the numerical solution at each step can be used to evaluate both sides of equation (27*b*) to form a residual. This residual can be monitored as the computation proceeds and should be of the same order as the accuracy achieved by the integrator.

To be specific, we differentiate equation (27*b*) and make the replacements

$$y_1(\xi) = S(\xi), y_2(\xi) = \omega(\xi), y_3(\xi) = \frac{d\omega}{d\xi}(\xi) = \omega'(\xi). \tag{31}$$

From equations (27*a*), (27*b*), and (31), we obtain a system of differential equations

$$H_1 \frac{dy_1}{d\xi} + H_2 \frac{dy_3}{d\xi} = H_3 \tag{32a}$$

$$\frac{dy_1}{d\xi} + Q_1 \frac{dy_3}{d\xi} = Q_2 \tag{32b}$$

$$\frac{dy_2}{d\xi} = y_3 \tag{32c}$$

where

$$H_1 = 1 - \frac{2}{9} nT\phi^n G^{n-1}(\xi, y_1, y_2, y_3) \tag{33a}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

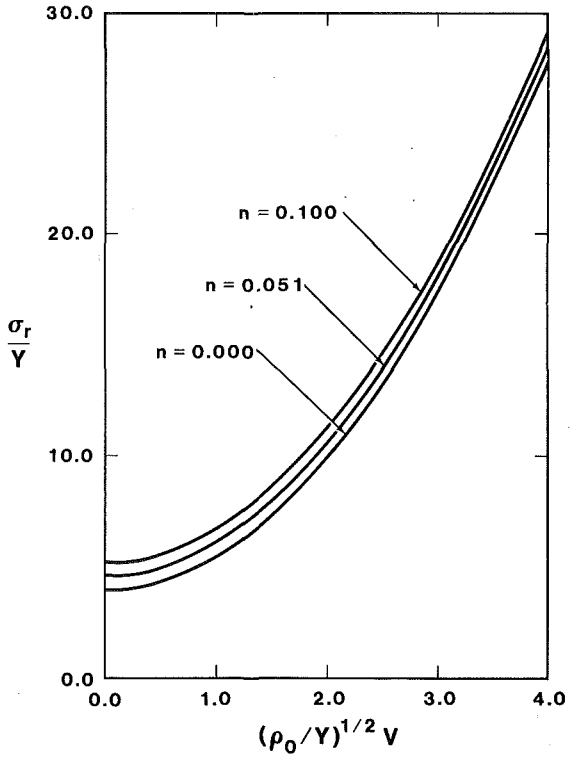


Fig. 5 Radial stress on the cavity surface versus cavity expansion velocity for an incompressible material

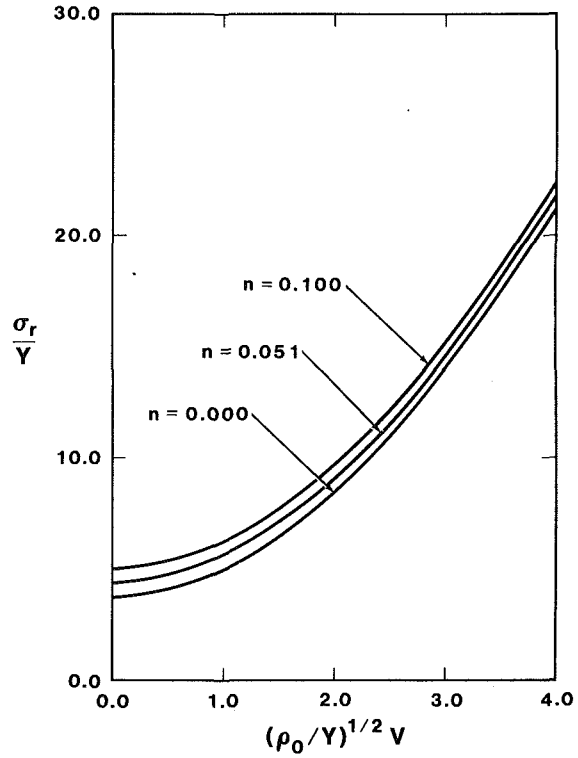


Fig. 6 Radial stress on the cavity surface versus cavity expansion velocity for a compressible material

$$H_2 = -\left[1 + \frac{2}{3} nT\phi^n G^{n-1}(\xi, y_1, y_2, y_3)\right]/y_3 \quad (33b)$$

$$H_3 = 2\left(\frac{y_3}{y_2} - \frac{1}{\xi}\right)\left[1 - \frac{2}{3} nT\phi^n G^{n-1}(\xi, y_1, y_2, y_3)\right] \quad (33c)$$

$$Q_1 = -\left(\frac{\lambda y_2^2}{\xi\, y_3}\right)^2 \quad (33d)$$

$$Q_2 = -3\left(\frac{2T\phi^n}{3}\right) G^n(\xi, y_1, y_2, y_3)/\xi \quad (33e)$$

$$G(\xi, y_1, y_2, y_3) = \ln\left(\frac{\xi^2 y_3}{y_2^2}\right) + \frac{y_1}{3}. \quad (33f)$$

Writing equation (32) in the standard form, we get

$$\frac{dy_1}{d\xi} = \frac{H_2 Q_2 - H_3 Q_1}{H_2 - H_1 Q_1} \quad (34a)$$

$$\frac{dy_2}{d\xi} = y_3 \quad (34b)$$

$$\frac{dy_3}{d\xi} = \frac{H_3 - H_1 Q_2}{H_2 - H_1 Q_1}. \quad (34c)$$

To start the numerical integration, values of $y_1 = S$, $y_2 = \omega$ and $y_3 = \omega'$ are needed at $\xi = 1$. $S(\xi = 1)$, denoted by S_1, is obtained from equation (30*a*). $y_2(\xi = 1) = \omega(\xi = 1)$ is obtained by solving the nonlinear algebraic equation formed by eliminating $\omega'(\xi = 1)$ from equations (28*b*), (28*c*) while $y_3(\xi = 1) = \omega'(\xi = 1)$ is computed from equation (28*b*) after $\omega(\xi = 1)$ is found. This nonlinear equation for the root $\omega(\xi = 1)$ is solved by SLATEC library (Cowell, 1984; Vandevender, 1984) code FZERO that uses a combination of the secant rule and bisection to compute the root. The working tolerance for all computations was set at 10^{-8} on a CDC machine that carries approximately 14 digits in the floating point mantissa.

While the integration of equation (34) proceeds normally for most of the range, the system gets progressively more difficult to integrate because $d\omega/d\xi = dy_2/d\xi$ becomes very large as $\omega = y_2 \to 0$. It is therefore desirable to use a state-of-the-art integrator that can monitor local error and adjust the integration step size accordingly. The SLATEC library code DEABM (Adams-Bashforth method) developed by Shampine and Watts (1980) is used because of the robust nature of the code. When the requested output step becomes too large (as signaled by DEABM), the problem is restarted at the last successful output point, and the output step is cut by a factor of 10. This process is repeated to refine the value of γ and obtain an estimate of γ with an accuracy on the order of 10^{-5}. For the cases considered, the residual formed from equation (27*b*) was always on the order of the requested accuracy.

Numerical Results and Discussion

We present graphical results from these cavity expansion analyses in Figs. 3 to 7. In Fig. 2 we compare the post-yield, power-law data-fit and the stress-strain data for 6061-T651 aluminum. The data in Fig. 2 were obtained in a large-strain compression test on a transverse specimen at a strain rate of $10^{-2}\, s^{-1}$. To show the relative effects of strain hardening, we also present numerical results for an elastic, perfectly plastic material and a material with more hardening than 6061-T651 aluminum (see Fig. 2). The power-law data-fit is given by equation (4*b*), and the material parameters are taken as $\rho_0 = 2{,}710$ kg/m^3, $E = 68.9$ GPa, $Y = 276$ MPa, and $\nu = 1/3$.

Figure 3 shows the elastic-plastic interface velocities c versus the cavity expansion velocity V for $n = 0.051$. The interface velocity for the incompressible model is given by equation (18) and computed numerically for the compressible model. We also computed c for $n = 0$ and 0.100, but no differences could be detected in the plotted results.

In Fig. 4, we compare radial stress profiles from the incompressible and the compressible models for $(\rho_0/Y)^{1/2} V = 2.70$ ($V = 862$ m/s) and $n = 0.051$. As previously discussed, radial

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

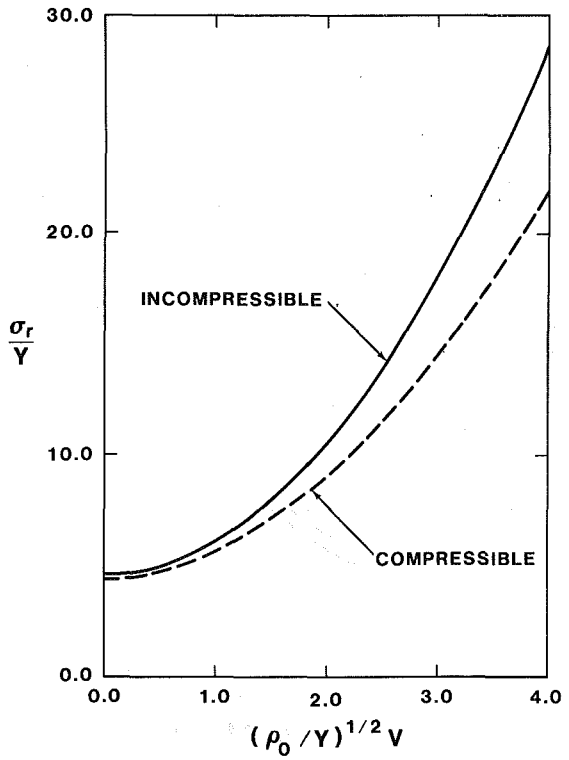


Fig. 7 Comparison of radial stresses on the cavity surfaces versus cavity expansion velocity for incompressible and compressible materials with n = 0.051

stress is continuous at the elastic-plastic interface for both models. Because radial stress is continuous at the elastic-plastic interface and the elastic stresses are independent of the strain-hardening exponent n, the radial stress in the plastic region at the interface must not depend on n. The numerical results show that radial stress has a weak dependence on n in the plastic region near the interface, and that continuity of radial stresses is achieved at the interface. Radial stress is also continuous at the elastic wave front for the compressible model, and there is no wave front for the incompressible model. As shown in Fig. 4, the incompressible model predicts larger stresses than the compressible model.

In the numerical procedure, we require continuity of radial stress and particle velocity at the elastic-plastic interface. To provide an additional check on the compressible solution, we calculated particle displacements at the interface to ensure material continuity. The displacements agreed to within the assigned accuracy. For the incompressible model, from equations (13) and (18) we found the displacement in the plastic region at the interface approximately equal to $Y/2E$ that is the elastic displacement at the interface from equation (17*a*).

As mentioned in the Introduction, we use results from these cavity expansion analyses to develop penetration equations. These penetration equations (Forrestal, Okajima, and Luk, 1988) require radial stress at the cavity surface versus cavity expansion velocity, so we present radial stresses at the cavity surface in Figs. 5, 6, and 7. Figures 5 and 6 show predictions for n = 0, 0.051, and 0.100 from the incompressible and compressible models, respectively. Figure 7 compares predictions from the incompressible and compressible models for n = 0.051. As expected, predictions from the two models diverge as cavity expansion velocity increases.

In summary, we developed models for the dynamic expansion of spherical cavities from zero initial radii for elastic-plastic, rate-independent materials with power-law strain hardening. Results from this study are used to formulate the penetration equations of Forrestal, Brar, and Luk (1991), who modeled 6061-T651 targets as an elastic, strain-hardening material.

Acknowledgments

This work was supported by the U.S. Department of Energy. The authors thank Dr. W. A. Kawahara for the data in Fig. 2.

References

Abramowitz, M., and Stegun, I. A., 1964, *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, 55.

Amos, D. E., 1988, "Evaluation of Integrals Related to the Debye Function," SAND88-2896, Sandia National Laboratories, Albuquerque, N. Mex.

Bishop, R. F., Hill, R., and Mott, N. F., 1945, "The Theory of Indentation and Hardness Tests," *The Proceedings of the Physical Society*, Vol. 57, Part 3, pp. 147–159.

Chadwick, P., 1959, "The Quasi-Static Expansion of a Spherical Cavity in Metals and Ideal Soils," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 12, Part 1, pp. 52–71.

Chakrabarty, J., 1987, *Theory of Plasticity*, McGraw-Hill, New York, p. 12.

Cowell, W. R., 1984, *Sources and Development of Mathematical Software*, Prentice-Hall, Englewood Cliffs, NJ.

Forrestal, M. J., and Luk, V. K., 1988, "Dynamic Spherical Cavity-Expansion in a Compressible Elastic-Plastic Solid," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 275–279.

Forrestal, M. J., Okajima, K., and Luk, V. K., 1988, "Penetration of 6061-T651 Aluminum Targets With Rigid Long Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 755–760.

Forrestal, M. J., Brar, S. N., and Luk, V. K., 1991, "Penetration of Strain-Hardening Targets with Rigid Spherical-Nose Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 7–10.

Hill, R., 1948, "A Theory of Earth Movement Near a Deep Underground Explosion," Memo No. 21-48, Armament Research Establishment, Fort Halstead, Kent, U.K.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Oxford Univ. Press, London.

Hill, R., 1980, "Cavitation and the Influence of Headshape in Attack of Thick Targets by Nondeforming Projectiles," *Journal of the Mechanics and Physics of Solids*, Vol. 28, pp. 249–263.

Hopkins, H. G., 1960, "Dynamic Expansion of Spherical Cavities in Metals," *Progress in Solid Mechanics*, Vol. 1, I. N. Sneddon, and R. Hill, eds., North-Holland, Amsterdam. Chapter 3.

Hunter, S. C., and Crozier, R. J. M., 1968, "Similarity Solution for the Rapid Uniform Expansion of a Spherical Cavity in a Compressible Elastic-Plastic Solid," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 21, Part 4, pp. 467–486.

Kawahara, W. A., 1986, "Compression Materials Testing at Low to Medium Strain Rates," Paper No. 86-WA/Mats-15, Presented at the Winter Annual Meeting, ASME, Anaheim, Calif., Dec. 7–12, 1986.

Mendelson, A., 1968, *Plasticity: Theory and Application*, Robert E. Krieger Publishing Co., p. 137.

Petzold, L., 1982, "A Description of DASSL: A Differential/Algebraic System Solver," SAND82-8637, Sandia National Laboratories, Albuquerque, N. Mex.

Shampine, L. F., and Watts, H. A., 1980, "DEPAC-Design of a User Oriented Package of ODE Solvers," SAND79-2374, Sandia National Laboratories, Albuquerque, N. Mex.

Stepelman, R., et al., eds., 1983, *IMACS Transactions on Scientific Computation*, Vol. 1, North-Holland, Amsterdam, p. 65.

Vandevender, W., 1984, "The SLATEC Mathematical Subprogram Library, Version 2.0," SAND84-0281, Sandia National Laboratories, Albuquerque, N. Mex.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

M. J. Forrestal
Sandia National Laboratories,
Albuquerque, NM 87185
Fellow ASME

N. S. Brar
University of Dayton Research Institute,
Dayton, OH 45469

V. K. Luk
Sandia National Laboratories,
Albuquerque, NM 87185
Mem. ASME

Penetration of Strain-Hardening Targets With Rigid Spherical-Nose Rods[1]

We developed engineering models that predict forces and penetration depth for long, rigid rods with spherical noses and rate-independent, strain-hardening targets. The spherical cavity expansion approximation simplified the target analysis, so we obtained closed-form penetration equations that showed the geometric and material scales. To verify our models, we conducted terminal-ballistic experiments with three projectile geometries made of maraging steel and 6061-T651 aluminum targets. The models predicted penetration depths that were in good agreement with the data for impact velocities between 0.3 and 1.0 km/s.

Introduction

In a recent paper, Forrestal, Okajima, and Luk (1988) developed penetration equations for long, rigid rods with spherical, ogival, and conical noses that penetrated rate-independent, elastic, perfectly plastic targets. The cavity expansion approximations (Bishop, Hill, and Mott, 1945) simplified the target analyses, so penetration equations were obtained in closed form. These models predicted penetration depths that were in good agreement with measurements for impact velocities between 0.4 and 1.4 km/s.

In this study, we modeled 6061-T651 aluminum targets as an elastic, strain-hardening, rate-independent material. We obtained compression stress-strain data to 100 percent true strain from samples taken from the target material. As shown in Fig. 1, a power-law data-fit closely approximated the post-yield stress-strain data. In a previous article (Forrestal, Okajima, and Luk, 1988) tensile tests were conducted with the target material, and the tensile specimen failed between true strains of 0.052 and 0.059. With this limited data, the authors assumed that the target material was elastic, perfectly plastic, and the flow stress was taken as the tensile failure stress. Therefore, the newly developed large-strain compression test (Kawahara, 1986) and power-law, strain-hardening model provided a more realistic description of the target material.

In addition to providing a realistic material description for the target, the penetration equations derived in this study showed the geometric and material scales. To verify our models, we conducted terminal-ballistic experiments with three projectile geometries made of maraging steel and 6061-T651 aluminum targets. The models predicted penetration depths that were in good agreement with measurements for impact velocities between 0.3 and 1.0 km/s.

Penetration Equations

Rigid, long-rod projectiles with spherical noses impact a uniform target at normal incidence with velocity V_0 and proceed to penetrate at rigid-body velocity, V_z. Much of the previous work by Forrestal, Okajima, and Luk (1988), who derived penetration equations for forces on the projectile nose and penetration depth, could be used in this study. Those equations employ the spherical cavity expansion approximation and assume that the target behaves as an elastic, perfectly plastic

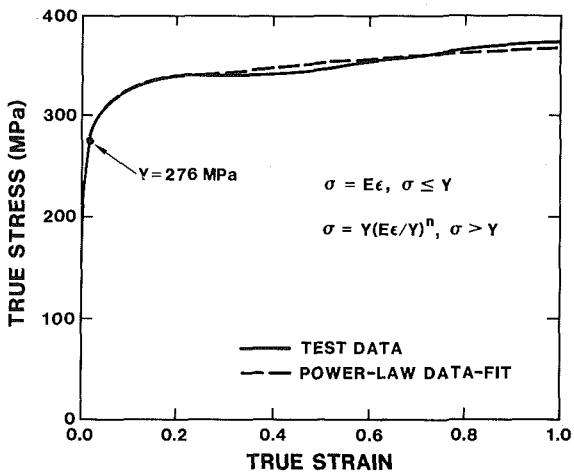


Fig. 1 Stress-strain data for the 6061-T651 aluminum target material and the power-law data-fit with E = 68.9 GPa, Y = 276 MPa, and n = 0.051

[1]Presented at the ASME Winter Annual Meeting, Symposium on "Computational Techniques for Contact, Impact, Penetration, and Perforation of Solids," San Francisco, Calif., Dec. 10–15, 1989.

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 3, 1989; final revision, Jan. 19, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

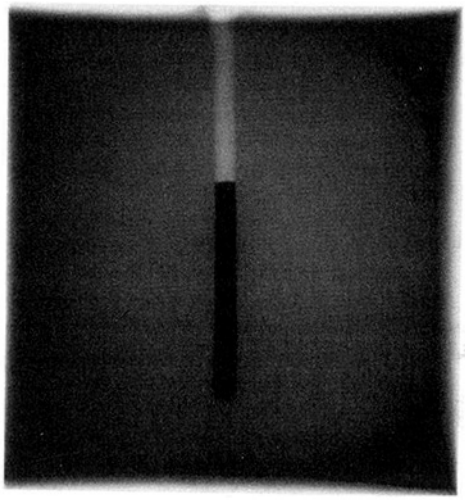


Fig. 2 Post-test, in-material X-ray photograph for the rod with $2a$ = 7.11 mm, L = 71.12 mm (shot number 1912)

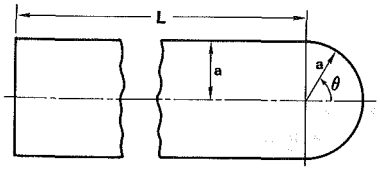


Fig. 3 Projectile geometry

material. In this study, we model the target as an elastic, strain-hardening, rate-independent material and arrange the penetration equations to show the geometric and material scales.

To obtain closed-form penetration equations, we use the spherical cavity expansion approximation (Bishop, Hill, and Mott, 1945). Normal stresses on the projectile nose are approximated by results from spherically symmetric, cavity-expansion analyses for an elastic, strain-hardening material (Luk, Forrestal, and Amos, 1991). In particular, normal stresses on the spherical nose are taken as the stresses at the cavity surface associated with the distributed expansion velocities at the nose-target interface caused by the projectile penetrating at rigid-body velocity V_z.

A post-test X-ray photograph (Fig. 2) shows that the projectile produces a tunnel in the target about the size of the shank diameter. As previously discussed (Forrestal, Okajima, and Luk, 1988), photomicrographs of the targets showed that a 5 to 15-μm film on the tunnel surface had undergone microstructural changes. From these observations, we assumed that during penetration a film of melted target material was between the nose and target. Thus, we take penetration resistance to consist of stresses normal to the nose and a tangential, frictional component from the melted film. We assume the tangential stress on the nose σ_t is proportional to the normal stress σ_n, so

$$\sigma_t = \mu \sigma_n \tag{1}$$

where μ is the sliding-friction coefficient.

Axial forces for spherical-nose rods (Forrestal, Okajima, and Luk, 1988) are given by

$$F_z = \pi a^2 \int_0^{\pi/2} \sigma_n(V_z, \theta)\left[\sin 2\theta + 2\mu \sin^2\theta\right] d\theta \tag{2}$$

where $\sigma_n(V_z, \theta)$ is the stress normal to the spherical nose and θ and a are defined in Fig. 3. The stress $\sigma_n(V_z, \theta)$ is approximated by results from spherically symmetric, cavity expansion analyses. For this cavity expansion problem (Luk, Forrestal, and Amos, 1991), a spherically symmetric cavity is expanded from

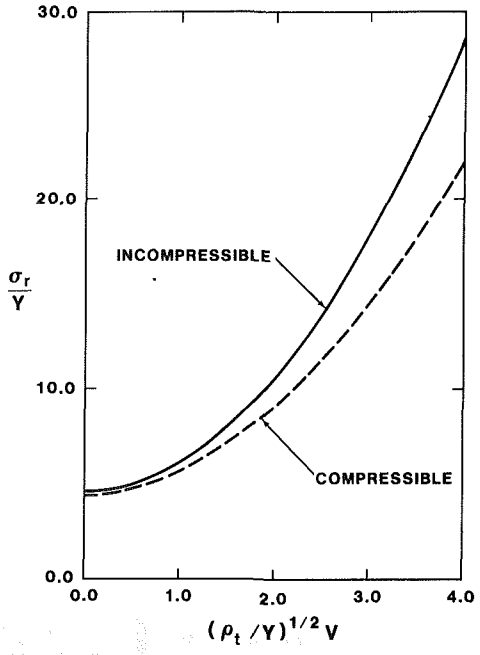


Fig. 4 Radial stress at the cavity surface versus cavity expansion velocity for an elastic, strain-hardening material with E = 68.9 GPa, ν = 1/3, Y = 276 MPa, n = 0.051, and ρ_t = 2,710 kg/m^3

zero initial radius at a constant velocity V. Figure 4 shows radial stress σ_r at the cavity surface versus the constant, cavity expansion velocity V for incompressible and compressible materials with the stress-strain behavior shown in Fig. 1. In Figs. 1 and 4, E, ν, Y, n, and ρ_t are Young's modulus, Poisson's ratio, yield stress, strain-hardening exponent, and density of the target material, respectively. The spherically symmetric, cavity expansion results in Fig. 4 can be expressed in the form

$$\sigma_r / Y = A_s + B_s\left[(\rho_t/Y)^{1/2} V\right]^2 \tag{3}$$

where A_s and B_s depend only on the target's material properties.

For an incompressible, elastic strain-hardening material (Luk, Forrestal, and Amos, 1991),

$$A_s = \frac{2}{3}\left[1 + \left(\frac{2E}{3Y}\right)^n I\right], \quad B_s = 3/2 \tag{4a}$$

$$I = \int_0^{1-(3Y/2E)} \frac{(-1n\, x)^n}{1-x}\, dx \tag{4b}$$

where I is evaluated numerically (Amos, 1988). For the material in Fig. 4, I = 4.539 and A_s = 4.595. To obtain closed-form penetration equations for the compressible model that show the geometric and material scales, we curve-fit the spherically symmetric, cavity expansion results in Fig. 4 with equation (3). In equation (3) and Fig. 4, A_s is the value of σ_r/Y for $V \to 0$, so only B_s is adjusted to fit the cavity expansion results. The compressible results in Fig. 4 are curve-fit accurately with A_s = 4.407 and B_s = 1.133.

From the geometry in Fig. 3, the target particle velocity at the nose-target interface caused by the projectile penetrating at rigid-body velocity V_z is

$$v(V_z, \theta) = V_z \cos\theta. \tag{4c}$$

We approximate the normal stress distribution on the spherical nose σ_n by replacing the spherically symmetric, cavity expansion velocity V in equation (3) with the particle velocity v given by equation (4c). Thus, normal stress distribution around the spherical nose is taken as

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Data summary for rods with $2a = 7.11$ mm and $L = 71.12$ mm

Shot Number	Mass (g)	V_0 (km/s)	α, β (degrees)	P (mm)	Target Length (mm)
2037	23.36	0.359	0.0, 1.2	21	90
1961	23.34	0.430	- , 0.2	32	89
1960	23.34	0.490	- , 0.7	40	82
1916	23.43	0.519	2.0, 2.8	41	114
1914	23.40	0.673	0.7, 1.0	73	178
1915	23.44	0.792	2.6, 1.8	85	178
2059	23.32	0.959	1.0, 0.5	109	198
1912	23.32	1.009	1.0, 0.2	129	216

Table 2 Data summary for rods with $2a = 5.08$ mm and $L = 71.12$ mm

Shot Number	Mass (g)	V_0 (km/s)	α, β (degrees)	P (mm)	Target Length (mm)
2041	11.76	0.253	0.2, 0.8	12	90
2040	11.72	0.357	0.4, 0.1	20	107
1936	11.75	0.425	1.5, 1.8	27	102
1931	11.76	0.457	1.4, 1.0	33	92
1928	11.74	0.599	0.8, 0.5	56	140
1932	11.82	0.740	0.2, 0.4	80	178
1933	11.85	0.866	0.6, 0.9	102	203
1935	11.74	0.980	0.6, 0.8	117	203

Table 3 Data summary for rods with $2a = 7.11$ mm and $L = 35.56$ mm

Shot Number	Mass (g)	V_0 (km/s)	α, β (degrees)	P (mm)	Target Length (mm)
2044	12.09	0.318	1.4, 2.8	12	90
1956	12.07	0.473	1.1, 1.9	23	102
1922	12.00	0.634	- , 0.9	38	140
1957	12.07	0.743	0.0, 2.3	47	89
1920	12.08	0.806	0.1, 0.8	56	127
1918	12.09	0.945	1.0, 1.1	68	152

$$\sigma_n(V_z, \theta)/Y = A_s + B_s[(\rho_t/Y)^{1/2} V_z \cos\theta]^2. \quad (5)$$

Substitution of equation (5) into equation (2) and integrating gives

$$F_z = \pi a^2 Y[A_s(1+\mu\pi/2) + B_s(\rho_t/Y)V_z^2(1+\mu\pi/4)/2]. \quad (6)$$

Final penetration depth is obtained from

$$m(dV_z/dt) = mV_z(dV_z/dz) = -F_z \quad (7)$$

where m is the projectile mass and z is penetration depth. The projectile mass is

$$m = \pi a^2 \rho_p (L + 2a/3) \quad (8)$$

where ρ_p is the projectile density and a and L are as defined in Fig. 3. Substitution of equations (6) and (8) into equation (7) and integrating gives

$$\frac{P}{(L+2a/3)} = \frac{(\rho_p/\rho_t)}{B_s(1+\mu\pi/4)} \ln\left[1 + \frac{1}{2}\left(\frac{B_s}{A_s}\right)\left(\frac{1+\mu\pi/4}{1+\mu\pi/2}\right)\left(\frac{\rho_t V_0^2}{Y}\right)\right] \quad (9)$$

where P is final penetration depth and V_0 is impact velocity. As previously mentioned, A_s and B_s depend only on the material properties of the target. Thus, the left side of equation (9) contains the geometric scale, and the right side contains the material scale and the effect of impact velocity.

Experiments

Targets. The targets were cut from 152-mm diameter (6.0-in.) 6061-T651 aluminum bars, and target length for each test is given in Tables 1, 2, and 3. In addition, axial and transverse specimens from this bar stock were machined for large strain, compression tests (Kawahara, 1986). Four specimens (two axial and two transverse) were tested at a strain rate of 10^{-2} s^{-1}. All stress-strain curves were nearly identical, and a test result for a transverse specimen is shown in Fig. 1. Maiden and Green (1966) performed compression tests on 6061-T6 aluminum for strain rates between 10^{-3} and 10^{3} s^{-1} and showed this material to be strain-rate independent over this range. However, Kawahara obtained true strains to 1.0, whereas Maiden and Green obtained true strains to 0.1. We used the data in Fig. 1 for our target-material model and assumed that the target material was strain-rate independent.

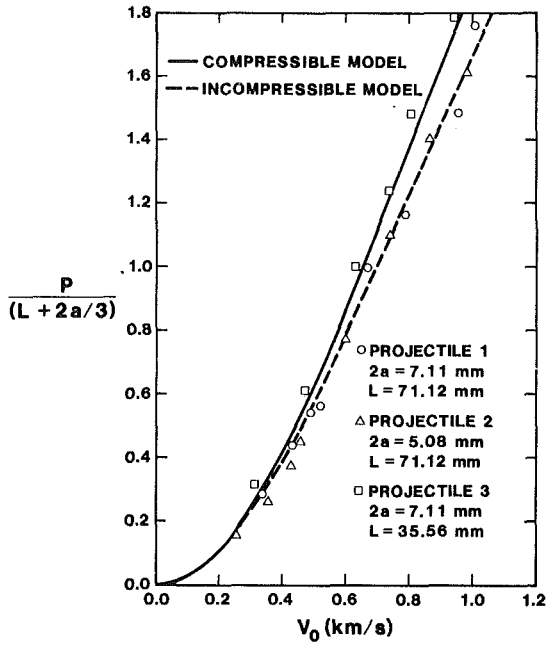


Fig. 5 Scaled depth of penetration and measurements

Projectiles. Projectiles (Fig. 3) were machined of T-200 maraging steel[2] with density $\rho_p = 8{,}000$ kg/m^3. The mass, length, and diameter for the projectiles are given in Tables 1, 2, and 3.

Terminal-Ballistic Experiments. A 20-mm, smooth-bore powder gun launched the rods shown in Fig. 3 to incident velocities between 0.3 and 1.0 km/s. The rods were encased in two-piece, serrated plastic sabots. Pusher plates were 15.7-mm diameter, 6.35-mm long titanium disks, and the obturator was made of the same material as the sabot. The sabot fit snugly in the gun barrel, while the slightly oversized obturator was forced into the barrel to provide a gas seal. Air drag stripped the sabot, pusher plate, and obturator from the rod prior to impact (Rosenberg and Forrestal, 1988).

Incident velocity was determined within one percent accuracy by the penetrator interrupting two continuously monitored laser beams, and incident pitch angle α and yaw angle β were measured with X-ray photographs. Definitions of pitch and yaw angles are those given in the book by Zukas et al. (1982). Post-test penetration depths were measured from in-material X-ray photographs; one of these photographs is shown in Fig. 2. Tables 1, 2, and 3 summarize the data for the experimental program.

[2]Vasco Pacific, 707 West Olympic Blvd., Montebello, CA 90640. T-200 maraging steel has a nominal yield stress of 200 ksi.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Comparison of Predicted and Measured Penetration Depths

Figure 5 shows predictions from the models that take the target material as incompressible and compressible. The target material properties were taken as $E = 68.9$ GPa, $\nu = 1/3$, $Y = 276$ MPa, $n = 0.051$, and $\rho_t = 2{,}710$ kg/m^3. Projectile geometries and masses ($\rho_p = 8{,}000$ kg/m^3) are given in Fig. 3 and Tables 1, 2, and 3. In addition, the models require the sliding-interface friction coefficient μ. Forrestal, Okajima, and Luk (1988) discuss experiments that determine μ at fast, sliding velocities. Unfortunately, values of μ are apparatus dependent and vary between 0.02 and 0.20. For Fig. 5, we used $\mu = 0.10$. Clearly, other approximations are made in the cavity expansion penetration models, so a selection of $\mu = 0.10$ may only compensate for other assumptions.

In Fig. 5, data for projectiles 1, 2, and 3 correspond to the geometries and data given in Tables 1, 2, and 3, respectively. The data trends show that projectiles 1 and 2 are in very close agreement for the scaled depth of penetration and that projectile 3 has a slightly larger scaled depth of penetration. Predictions using the incompressible and compressible models are in reasonably close agreement.

Summary and Conclusions

We developed engineering models that predicted penetration depth for long, rigid rods with spherical noses and rate-independent, strain-hardening targets. The spherical, cavity expansion approximation simplified the target analysis, so we obtained closed-form penetration equations that showed the geometric and material scales. Terminal-ballistic experiments with three projectile geometries verified the geometric scale. Any improvements in modeling will require new experimental methods that quantify the sliding frictional resistance between the projectile and target.

Acknowledgments

This work was supported by the U.S. Department of Energy. The authors thank Mr. Mark Laber for assisting with the terminal-ballistic experiments.

References

Amos, D. E., 1988, "Evaluation of Integrals Related to the Debye Function," SAND88-2896, Sandia National Laboratories, Albuquerque, N. Mex.

Bishop, R. F., Hill, R., and Mott, N. F., 1945, "The Theory of Indentation and Hardness Tests," *The Proceedings of the Physical Society*, Vol. 57, Part 3, pp. 147–159.

Forrestal, M. J., Okajima, K., and Luk, V. K., 1988, "Penetration of 6061-T651 Aluminum Targets with Rigid Long Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 4, pp. 755–760.

Kawahara, W. A., 1986, "Compression Materials Testing at Low to Medium Strain Rates," Paper No. 86-WA/Mats-15, presented at the ASME Winter Annual Meeting, Anaheim, Calif.

Luk, V. K., Forrestal, M. J., and Amos, D. E., 1989, "Dynamic Spherical Cavity Expansion of Strain-Hardening Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 58, pp. 1–6.

Maiden, C. J., and Green, S. J., 1966, "Compressive Strain-Rate Tests on Six Selected Materials at Strain Rates from 10^{-3} to 10^4 in./in./sec," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 496–504.

Rosenberg, Z., and Forrestal, M. J., 1988, "Perforation of Aluminum Plates with Conical-Nosed Rods—Additional Data and Discussion," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 236–238.

Zukas, J. A., et al., 1982, *Impact Dynamics*, John Wiley and Sons, New York, p. 206.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Yih-Hsing Pao
Fellow ASME.

Tsung-Tsong Wu[2]

Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853

Udo Gamer
Vienna Technological University.

Acoustoelastic Birefringences in Plastically Deformed Solids: Part I—Theory[1]

A modified theory of acoustoelasticity is proposed to include the effect of small plastic deformation on propagational speeds of elastic waves in orthotropic media. In the constitutive equation for the incremental stresses of a predeformed body, the elastic moduli are modified by elastic deformation and plastic strains. The latter is related to the texture of the material. Several formulae for the acoustoelastic birefringences (differences of two shear wave speeds polarized at perpendicular directions) are derived. They can be applied to measure residual stresses and plastic strains by ultrasonic techniques.

Introduction

In 1959, Benson and Raelson (1959) observed that the elastic wave speeds are changed by stresses in a solid and called the phenomenon the acoustoelastic effect. Since then, there has been considerable progress made on the theory and experiments of acoustoelasticity. The difference of the speeds of two shear waves polarized at two perpendicular directions are called the acoustoelastic birefringence. In analogous to photoelasticity, the birefringence was found to be proportional to the difference of two principal stresses in a plane specimen (Benson and Raelson, 1959).

A general theory of acoustoelasticity was proposed by Toupin and Bernstein (1961), and Thurston and Brugger (1964). When the theory was applied later to measure residual stresses by ultrasonic techniques, however, it encountered some difficulty (Crecraft, 1967). One reason is that the grains of polycrystalline materials, such as structural steel and aluminum, are oriented toward a preferred direction owing to manufacturing processes. The preferred orientations of crystalline grains, known also as the texture effect, induce anisotropy in the mechanical property of the materials. To account for the inherent anisotropy, several theories of acoustoelasticity were proposed later (Tokuoka and Saito, 1969; Iwashimizu and Kubomura, 1973; King and Fortunko, 1983; Pao and Gamer, 1985a).

A common assumption for the aforementioned theories of acoustoelasticity is that the material be hyperelastic, that is, there exists a strain energy function from which the stress-strain relation can be deduced. Hence, for a plastically deformed body, these theories can not be applied without modification. A review of the theories and experiments on the changes of elastic wave speeds in a predeformed body was made by Pao, Sachse, and Fukuoka (1984).

The preferred orientations of a given material are changed very little by elastic deformations of the material. Hence, the texture effect may be treated as an inherent anisotropy of the material. However, in the case of a body undergoing one or several cycles of elastoplastic loading and unloading, the interaction between the plastic deformation and preferred orientation can no longer be neglected. The induced change of preferred orientation by plastic deformation alters the inherent anisotropy, which in turn induces change in acoustoelastic birefringence. Such a texture-induced anisotropy is often as large as the anisotropy induced by the residual stresses in metals.

Experimental investigations of the effect of plastic deformation on the acoustoelastic response of several materials have been reported (Johnson, 1981b, Toda et al., 1982; Hirao and Pao, 1985; Wong and Johnson, 1985), and a few modified theories of acoustoelasticity were proposed to incorporate the plastic deformation effect (Johnson, 1981a, 1983; Kobayashi, 1987). The theory presented in this paper is an extension of that proposed by Pao and Gamer (Pao and Gamer, 1983; Pao, 1985b), who incorporated the effects of texture changes caused by plastic deformation into the constitutive equations of a modified hyperelastic theory. Based on the theory, new acoustoelastic birefringent formulae are derived. The birefringences are dependent on the difference of principal stresses as well as that of principal plastic strains.

In addition, experiments with steel specimens in uniaxial loading and bending were conducted to determine the plastic strain effect on the change of acoustoelastic birefringence. These are reported in Part II of this paper.

[1]Report No. 6291, Material Science Center, Cornell University.

[2]Now at the Institute of Applied Mechanics, National Taiwan University, Taipei, Taiwan.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, Dallas, Texas, Nov. 25–30, 1990.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, the Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 30, 1988; final revision, Nov. 14, 1989.

Paper No. 90-WA/APM-41.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

I Equation of Motion of a Predeformed Body

In the theory of elasticity, the deformation of an elastic body is described in terms of two states (configurations), the natural state (undeformed state, free of stress and strain) and the current state (deformed state). In the hyperelastic theory of acoustoelasticity, the deformed state is divided into two parts: the *initial state* (predeformed state) and the *final state* (current state). As for measuring residual stresses, which are generated by the interlocking of inhomogeneous permanent strains inside the body, another state called the *relaxed state* was introduced (Pao et al., 1984).

To fix ideas, the four states are illustrated by an experiment with a beam in plastic bending. A straight metal beam received in a laboratory is assumed to be in its *natural state*, the stress and strain at a typical point "O," say a point at the central section of the beam, being zero (Fig. 1). The beam is then loaded in a universal testing machine, first in the hyperelastic range (OY), then passing the yield point Y to the plastic range (YB). After total unloading, residual stresses remain in the beam (Point A) because of the inhomogeneous plastic deformation created by the plastic bending. The beam is then sectioned in both directions of depth and length (cut into grids) to relax the residual stresses. The sectioned beam after unloading (Point R) is in the relaxed state. As will be stated more clearly in Section III, the relaxed state is characterized by a new set of elastic constants which are different from those in the natural state, and the deformation from the relaxed state to the initial state (point A) is assumed to be hyperelastic.

During the loading-unloading process, an ultrasonic experiment is performed to measure the applied stress at point C or residual stress at point A. The body is in its initial state before the ultrasonic wave is applied (Point A, C, or any point on OCYBA curve), and in the final state after the ultrasonic wave is superposed on the predeformed body (the dynamic final state is not shown in the figure).

In many applications, the natural state of a specimen being tested is not known. For example, the straight beam received in the laboratory mentioned previously might have been rolled or drawn from a large metal block (an ingot). Often, one is asked to determine the residual stresses in the straight beam as received. For such a case, the beam can no longer be considered free of stress and strain and is assumed to be in the state with initial stress for ultrasonic tests. The natural state (the ingot) in this case is unknown to us, but to a certain degree, the relaxed state always can be found by cutting the beam.

In this paper, four position vectors $\boldsymbol{\xi}$, $\mathbf{X}$, $\mathbf{x}$, and $\mathbf{A}$ are used to denote the positions of a material point $\mathbf{Q}$ at four different states, the natural state ($\boldsymbol{\xi}$), initial state ($\mathbf{X}$), final state ($\mathbf{x}$), and relaxed state ($\mathbf{A}$). They are all referred to the common, fixed Cartesian coordinate system (Fig. 2). For purpose of distinguishing the two-point tensorial quantities in four states, different indices are used for the tensorial components. They are $\xi[\alpha, \beta, \gamma, \ldots]$, $\mathbf{X}[I, J, K, \ldots]$, $\mathbf{x}[i, j, k, \ldots]$, and $\mathbf{A}[R, S, T, \ldots]$.

From the natural state to the initial state, the deformation is assumed to be static, and the displacement of a material point is denoted by $\mathbf{u}^i$. From the natural to the final state, the displacement vector is $\mathbf{u}^f$ which depends on time. They are related to the position vectors by

$$\tilde{\mathbf{u}}^i(\boldsymbol{\xi}) = \tilde{\mathbf{X}}(\boldsymbol{\xi}) - \boldsymbol{\xi}, \quad \tilde{\mathbf{u}}^f(\boldsymbol{\xi},t) = \tilde{\mathbf{x}}(\boldsymbol{\xi},t) - \boldsymbol{\xi}, \tag{1}$$

where

$$\mathbf{X} = \tilde{\mathbf{X}}(\boldsymbol{\xi}), \quad \mathbf{x} = \tilde{\mathbf{x}}(\boldsymbol{\xi},t) = \hat{\mathbf{x}}(\mathbf{X},t). \tag{2}$$

The $\tilde{\mathbf{X}}(\boldsymbol{\xi})$ and $\tilde{\mathbf{x}}(\boldsymbol{\xi}, t)$ are the one-to-one mappings from the natural configuration to the initial and final configurations, respectively. The deformation from the initial to the final configuration is characterized by the mapping $\hat{\mathbf{x}}(\mathbf{X}, t)$.

The superposed ultrasonic wave $\mathbf{u}$, which is the difference between the initial state ($\mathbf{X}$) and the final state ($\mathbf{x}$), is given as

$$\mathbf{u} = \tilde{\mathbf{u}}(\boldsymbol{\xi},t) = \tilde{\mathbf{x}} - \tilde{\mathbf{X}} = \tilde{\mathbf{u}}^f - \tilde{\mathbf{u}}^i, \tag{3}$$

or

$$\mathbf{u} = \hat{\mathbf{u}}(\mathbf{X},t) = \hat{\mathbf{x}}(\mathbf{X},t) - \mathbf{X}. \tag{4}$$

The superscripts i and f are used, whenever necessary, to designate a physical quantity in the initial and final states, respectively. The tilde and caret are used to denote functions that are dependent on the natural ($\boldsymbol{\xi}$) and initial ($\mathbf{X}$) coordinates, respectively. A change from the initial to the final state, sometimes is referred to the incremental state, and a quantity in the incremental state is written without any superscript designation.

The Cartesian components of the Lagrangian strain tensors at the initial and final states when referred to the natural coordinates are defined as

$$E^i_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial \tilde{X}_K}{\partial \xi_\alpha}\frac{\partial \tilde{X}_K}{\partial \xi_\beta} - \delta_{\alpha\beta}\right),$$
$$E^f_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial \tilde{x}_k}{\partial \xi_\alpha}\frac{\partial \tilde{x}_k}{\partial \xi_\beta} - \delta_{\alpha\beta}\right). \tag{5}$$

Similarly, the incremental strain tensor $\mathbf{E}$, which is the Lagrangian strain, at the final state referred to the initial configuration is

$$E_{IJ} = \frac{1}{2}\left(\frac{\partial \hat{x}_k}{\partial X_I}\frac{\partial \hat{x}_k}{\partial X_J} - \delta_{IJ}\right). \tag{6}$$

It could also be deduced from the difference of $\mathbf{E}^f$ and $\mathbf{E}^i$,

$$E_{\alpha\beta} = E^f_{\alpha\beta} - E^i_{\alpha\beta}. \tag{7}$$

In this paper, σ^f_{ij} denote the Cauchy stress at the final state. It may be transformed to the second Piola-Kirchhoff stress (Jaunzemis, 1967), $T^f_{\alpha\beta}$, in the natural configuration, or that in the initial configuration, T^f_{IJ}. Similarly, the Cauchy stress in the initial state and the corresponding second Piola-Kirchhoff stress referred to the natural configuration are denoted by σ^i_{IJ} and $T^i_{\alpha\beta}$, respectively. As mentioned before, the indices of the tensorial components of these stresses are consistent with the configuration to which they are referred. In the sequel, the second Piola-Kirchhoff stress is often abbreviated as the Kirchhoff stress.

The equation of motion of a predeformed body can be written in terms of initial coordinates as (Tokuoka and Iwashimizu, 1968)

$$\frac{\partial}{\partial X_J}\left(T_{IJ} + \frac{\partial \hat{u}_I}{\partial X_K}\sigma^i_{JK}\right) = \rho^i \frac{\partial^2 \hat{u}_I}{\partial t^2}. \tag{8}$$

In this equation, $\mathbf{u}$ is the incremental displacement caused by the superimposed dynamic disturbance as defined in equation (4), and σ^i_{JK} is the initial stress in the predeformed body. The incremental second Piola-Kirchhoff stress (T_{IJ}) is defined as

$$\mathbf{T} = \mathbf{T}^f - \boldsymbol{\sigma}^i, \quad T_{IJ}(\mathbf{X},t) = T^f_{IJ}(\mathbf{X},t) - \sigma^i_{IJ}(\mathbf{X}). \tag{9}$$

It is important to note that this equation of motion for the small incremental motion (equation (8)) is valid for a general predeformation, infinitesimal or finite, elastic or inelastic, because no constitutive relation has been invoked in the derivations. The key point in the completion of theory is the finding of a constitutive equation for T_{IJ} during the propagation of acoustic waves in an elastoplastically deformed body.

II Constitutive Equations for Hyperelastic Deformation

For acoustoelastic experiments, the initial stress, $\boldsymbol{\sigma}^i$, in equation (8) is unknown. Its relation to the initial deformation, however, need not be specified at this stage. Since the incremental motion resulting from superimposed ultrasonic waves

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

is small, the incremental deformation is assumed to be elastic. To obtain the constitutive relation for the incremental Kirchhoff stress, which is another unknown in equation (8), we need to know the relation between T_{IJ} and $\partial \hat{u}_I/\partial X_J$ for a given material. The constitutive property, however, is very complicated because it is affected by material properties as well as initial elastic or elastoplastic deformation. The constitutive equations for the case of initial elastoplastic deformation will be discussed in Section III.

As a basis for the subsequent derivations and discussions of constitutive equations for elastoplastic deformation, we repeat briefly the derivation of the constitutive equation for hyperelastic deformation (Pao et al., 1984). For a hyperelastic body with isentropic deformation, the Kirchhoff stresses in the initial and final states are derivable from a strain energy function $W(\mathbf{E})$ (strain energy per unit volume)

$$T^{i,f}_{\alpha\beta} = \frac{\partial W(E^{i,f}_{\alpha\beta})}{\partial E^{i,f}_{\alpha\beta}}, \tag{10}$$

where i, f represent the initial and final states, respectively.

The $\mathbf{T}^f$ can also be expanded about the initial strain ($\mathbf{E}^i$) as

$$T^f_{\alpha\beta} = T^i_{\alpha\beta} + \left(\frac{\partial T^f_{\alpha\beta}}{\partial E^f_{\gamma\delta}}\right)^i (E^f_{\gamma\delta} - E^i_{\gamma\delta}) + \ldots, \tag{11}$$

and when the higher-order terms are omitted, the difference of $\mathbf{T}^f$ and $\mathbf{T}^i$ referred to the natural state can be expressed as

$$T_{\alpha\beta} = T^f_{\alpha\beta} - T^i_{\alpha\beta} = \left(\frac{\partial T^f_{\alpha\beta}}{\partial E^f_{\gamma\delta}}\right)^i E_{\gamma\delta}, \tag{12}$$

where the incremental strain $E_{\gamma\delta}$ is defined in equation (7).

The coefficients $\partial \mathbf{T}^f/\partial \mathbf{E}^f$ in equation (12) can thus be obtained by expanding the strain energy function $W(\mathbf{E}^f)$ into a power series of $\mathbf{E}^f$ and replacing all $\mathbf{E}^f$ in $\partial W/\partial \mathbf{E}^f$ by $\mathbf{E}^i$. After truncating the series into cubic polynomials, we obtain an approximate value

$$\left(\frac{\partial T^f_{\alpha\beta}}{\partial E^f_{\gamma\delta}}\right)^i = c^o_{\alpha\beta\gamma\delta} + c^o_{\alpha\beta\gamma\delta\epsilon\eta} E^i_{\epsilon\eta}. \tag{13}$$

The coefficients of expansion, $c^o_{\alpha\beta\gamma\delta}$ and $c^o_{\alpha\beta\gamma\delta\epsilon\eta}$, are the second and third-order elastic moduli of the material, all being measured in the natural state of the body.

By applying the relation between the two incremental Kirchhoff stresses T_{IJ} and $T_{\alpha\beta}$,

$$T_{IJ} = \frac{\rho^i}{\rho^o} \frac{\partial \tilde{X}_I}{\partial \xi_\alpha} \frac{\partial \tilde{X}_J}{\partial \beta_\beta} T_{\alpha\beta}, \tag{14}$$

and the relation between $E_{\gamma\delta}$ and E_{KL},

$$E_{\gamma\delta} = \frac{\partial \tilde{X}_K}{\partial \xi_\gamma} \frac{\partial \tilde{X}_L}{\partial \xi_\delta} E_{KL}, \tag{15}$$

we find a constitutive relation for the incremental Kirchhoff stress tensor at initial coordinates

$$T_{IJ} = C^o_{IJKL} E_{KL} \cong C^o_{IJKL} \frac{\partial u_L}{\partial X_K}. \tag{16}$$

In equation (16), the incremental elastic coefficient C^o_{IJKL}, is given by (Toupin and Bernstein, 1961)

$$C^o_{IJKL} = \frac{\rho^i}{\rho^o} \frac{\partial \tilde{X}_I}{\partial \xi_\alpha} \frac{\partial \tilde{X}_J}{\partial \xi_\beta} \frac{\partial \tilde{X}_K}{\partial \xi_\gamma} \frac{\partial \tilde{X}_L}{\partial \xi_\delta} \left(\frac{\partial T^f_{\alpha\beta}}{\partial E^f_{\gamma\delta}}\right)^i. \tag{17}$$

It is seen from equation (17) that the C^o_{IJKL} is influenced by the geometry of the initial deformation gradient as well as the material dependent function $\partial T^f_{\alpha\beta}/\partial E^f_{\gamma\delta}$.

Equation (17) can be further simplified by substituting the deformation gradients with the displacement gradients

$$\frac{\partial \tilde{X}_I}{\partial \xi_\alpha} = \delta_{I\alpha} + \frac{\partial u^i_I}{\partial \xi_\alpha}, \tag{18}$$

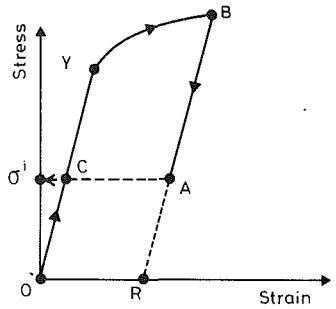


Fig. 1 The stress-strain curve of a typical point in a plastically predeformed body

where $\delta_{I\alpha}$ is the shift factor. On retaining the linear terms in the expansions of equation (17) and making use of equation (13), the incremental elastic coefficient becomes

$$C^o_{IJKL} = c^o_{IJKL}\left(1 - \frac{\partial u^i_M}{\partial X_M}\right) + c^o_{MJKL}\frac{\partial u^i_I}{\partial X_M} + c^o_{IMKL}\frac{\partial u^i_J}{\partial X_M} + c^o_{IJML}\frac{\partial u^i_K}{\partial X_M} + c^o_{IJKM}\frac{\partial u^i_L}{\partial X_M} + c^o_{IJKLMN}\frac{\partial u^i_N}{\partial X_M}, \tag{19}$$

where $c^o_{IJKL} = \delta_{I\alpha}\delta_{J\beta}\delta_{K\gamma}\delta_{L\delta}c^o_{\alpha\beta\gamma\delta}$. This linearization implies that the initial strain is infinitesimal, $E^i_{IJ} \cong e^i_{IJ} \equiv 1/2\,(\partial u^i_I/\partial X_J + \partial u^i_J/\partial X_I)$.

For a slightly orthotropic body, the second-order and third-order elastic constants in the incremental elastic coefficient (equation (19)) can be approximated by (Iwashimizu and Kubomura, 1973),

$$c^o_{IJKL} = c'_{IJKL} + c''_{IJKL}, \qquad c^o_{IJKLMN} = c'_{IJKLMN} + c''_{IJKLMN}, \tag{20}$$

where

$$c'_{IJKL} = \lambda\delta_{IJ}\delta_{KL} + \mu(\delta_{IK}\delta_{JL} + \delta_{IL}\delta_{JK}), \tag{21}$$

and

$$|c''_{IJKL}/\mu| \ll 1, \quad |c''_{IJKLMN}/\nu_p| \ll 1.$$

The c'_{IJKL} and c'_{IJKLMN} are the second and third-order elastic constants of an isotropic body, λ and μ are the Lamè constants, and ν_p is a third-order elastic constant. Substitution of equation (19) in equation (16) and then in equation (8) completes the equation of motion for a predeformed hyperelastic body. Note that in the theory, the elastic coefficient $\mathbf{c}^o$ at the natural state is changed to $\mathbf{C}^o$, which depends on the initial displacement gradients. The elastic moduli of the material are thus changed by the predeformation.

III Constitutive Equations for Elastoplastic Deformation

To incorporate the plastic deformation into the theory of acoustoelasticity, we make the following additional assumptions:

(*a*) The predeformed body (initial state) is in equilibrium and the anelastic effects resulting from plastic deformations are neglected.

(*b*) The plastic (permanent) strains are small. Although the change of crystalline anisotropy (the texture effect) may be a complicated function of plastic deformation, or even should be treated as a new state variable, a linear dependence on the plastic strains is assumed.

(*c*) The superposed dynamic motions are small and do not cause additional work hardening of the material.

(*d*) The unloading as well as the reloading from the relaxed state are assumed to be hyperelastic. A hyperelastic constitutive

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

relation with new elastic moduli can thus be deduced for each relaxed state.

To make the concept clear, we illustrate the changes of elastic constants again by the stress-strain curve of a material in uniaxial test or a typical point in a plastically bent beam (Fig. 1). The solid line in Fig. 1 shows the elastoplastic loading history of the specimen. So long as the loading remains in the hyperelastic range OY, the elastic constants remain the same as those in the natural state (point O). As the loading passes the yield limit Y and then to point A, however, the preferred orientations of the material are changed by the plastic deformation. Hence, the elastic constants in this state (initial state, point A) are different from those in the natural state (point O). Instead of following the complicated elastoplastic loading history (OYBA), the residual stresses at point A relative to the natural state O may be regarded as the applied stresses at A relative to the relaxed state R. The material is characterized by a new set of elastic constants in the relaxed state (point R) and is hyperelastic from the relaxed state to the initial state A. From the natural state O to the relaxed state R, we assume a homogeneous plastic deformation.

Over the entire body, the plastic deformation, which induces residual stresses, is nonhomogeneous. The constitutive property of the material, however, is defined within a small neighborhood around a material point, and hence, it is reasonable to assume that the deformation is homogeneous locally at the relaxed state.

In the deformation, a material point at the natural state, $Q(\xi)$, is mapped to the point $Q^r(\mathbf{A})$ at the relaxed state, and a point $S(\xi + d\xi)$ in the vicinity of $Q(\xi)$ is transformed into a point $S^r(\mathbf{A} + d\mathbf{A})$ in the vicinity of $Q^r(\mathbf{A})$, where

$$\mathbf{dA} = \frac{\partial \mathbf{A}}{\partial \xi}\,\mathbf{d\xi}, \quad dA_R = \frac{\partial A_R}{\partial \xi_\alpha}\,d\xi_\alpha. \tag{22}$$

For homogeneous deformation, $\partial A/\partial \xi$ is constant. By this transformation, parallel straight lines are mapped into parallel straight lines, and a cube is mapped into a rectangular parallelepiped. If the displacement of the point Q (to Q^r) is denoted by $\mathbf{u}^p$ and $\mathbf{A} = \xi + \mathbf{u}^p$, we have

$$\frac{\partial A_R}{\partial \xi_\alpha} = \delta_{R\alpha} + \frac{\partial u^p_R}{\partial \xi_\alpha}. \tag{23}$$

Following the assumptions (a) and (b), we assume that the second-order elastic constants $\mathbf{c}^r$ in the relaxed state are related to those in the natural state $\mathbf{c}^o$ as

$$c^r_{RSTU} = c^o_{\alpha\beta\gamma\delta}\delta_{\alpha R}\delta_{\beta S}\delta_{\gamma T}\delta_{\delta U} + p^o_{\alpha\beta\gamma\delta}\left(\frac{\partial A_R}{\partial \xi_\alpha}\frac{\partial A_S}{\partial \xi_\beta}\frac{\partial A_T}{\partial \xi_\gamma}\frac{\partial A_U}{\partial \xi_\delta} - \delta_{\alpha R}\delta_{\beta S}\delta_{\gamma T}\delta_{\delta U}\right). \tag{24}$$

In these equations, the subscripts "R, S, T, U, V, W" are used to represent the components of a tensor in the relaxed state. The first term is the shifting of natural elastic constants to the relaxed state by the relations $c^o_{\alpha\beta\gamma\delta} = c_{RSTU}\,\delta_{R\alpha}\delta_{S\beta}\,\delta_{T\gamma}\,\delta_{U\delta} = c^o_{IJKL}\,\delta_{I\alpha}\,\delta_{J\beta}\,\delta_{K\gamma}\,\delta_{L\delta}$. The second term represents the effect of plastic deformation on the elastic moduli, $\mathbf{c}^o$, at the natural state. The texture constants $p^o_{\alpha\beta\gamma\delta}$ are introduced to account for the change of preferred orientations of the crystallines. The first part of the second term is analogous to the change of $\mathbf{c}^o$ to $\mathbf{C}^o$ in equation (17) and the second part with the minus sign implies that the rigid body translation in the plastic deformation causes no texture change.

Substituting equation (23) into equation (24), and neglecting nonlinear terms in the products, we express $(\mathbf{c}^r)$ as

$$c^r_{RSTU} = c^o_{RSTU} + c^p_{RSTU}, \tag{25}$$

where

$$c^p_{RSTU} = p^o_{VSTU}\,e^p_{RV} + p^o_{RVTU}\,e^p_{SV} + p^o_{RSVU}\,e^p_{TV} + p^o_{RSTV}\,e^p_{UV}. \tag{26}$$

In the derivation, the approximation $\partial \mathbf{u}^p/\partial \xi \cong \partial \mathbf{u}^p/\partial \mathbf{A}$ has been made, and the rotational part of the infinitesimal displacement gradient vanishes because the plastic deformation is homogeneous. The coefficients c^p_{RSTU}, which represent the change of the elastic constants due to a change of texture in metals, are linear functions of the plastic strains $\mathbf{e}^p$. This approximation also implies that $|\mathbf{c}^p|$ are much less than $|\mathbf{c}^o|$.

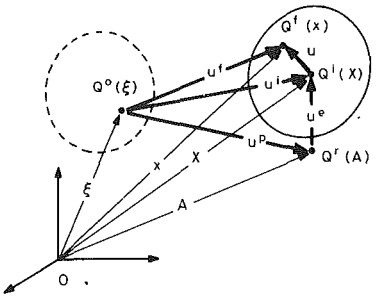


Fig. 2 Positions and displacements of a material point at the natural state $Q^o(\xi)$, initial state $Q^i(X)$, final state $Q^f(x)$, and relaxed state $Q^r(A)$

The deformation at the initial state is now regarded as the superposition of the deformation from the natural to the relaxed state and that from the relaxed to the initial state (Fig. 2)

$$\mathbf{A} = \xi + \mathbf{u}^p, \quad \mathbf{X} = \mathbf{A} + \mathbf{u}^e. \tag{27}$$

The total strain in the initial state may now be expressed as

$$\mathbf{E}^i = \mathbf{E}^p + \mathbf{E}^e; \quad E^i_{IJ} = E^p_{IJ} + E^e_{IJ}, \tag{28}$$

the first part of the deformation is purely plastic (unrecoverable) and the second part purely elastic.

Given the new set of elastic constants c^r_{RSTU} (equation (25)) at the relaxed state (point R in Fig. 1), the modified incremental elastic coefficients in the initial state (point A) relative to the relaxed state is expressed as

$$C^r_{IJKL} = \frac{\rho^i}{\rho^o}\frac{\partial X_I}{\partial A_R}\frac{\partial X_J}{\partial A_S}\frac{\partial X_K}{\partial A_T}\frac{\partial X_L}{\partial A_U}\,(c^r_{RSTU} + c^o_{RSTUVW}E^e_{VW}). \tag{29}$$

This relation is analogous to equation (17) for the hyperelastic material. In the equation, we have also assumed that the third-order elastic coefficients are not changed by the plastic strains, that is, $c^r_{RSTUVW} \cong c^o_{RSTUVW}$, and $\mathbf{E}^e$ is the elastic part of the initial Lagrangian strain $\mathbf{E}^i$. If the entire deformation is hyperelastic, the relaxed coordinates $\mathbf{A}$ coincide with the natural coordinates ξ, and the C^r_{IJKL} in equation (29) reduce to C^o_{IJKL} in equation (17).

By retaining only the linear terms in elastic displacement gradients, and noting that c^p_{IJKL} is small compared with c^o_{IJKL}, equation (29) is reduced to

$$C^r_{IJKL} = c^p_{IJKL} + c^o_{IJKL}\left(1 - \frac{\partial u^e_M}{\partial X_M}\right) + c^o_{MJKL}\frac{\partial u^e_J}{\partial X_M} + c^o_{IMKL}\frac{\partial u^e_J}{\partial X_M} + c^o_{IJML}\frac{\partial u^e_K}{\partial X_M} + c^o_{IJKM}\frac{\partial u^e_L}{\partial X_M} + c^o_{IJKLMN}\frac{\partial u^e_M}{\partial X_N}. \tag{30}$$

The constitutive relation for the incremental stress and strain of an elastoplastically predeformed body becomes

$$T_{IJ} = C^r_{IJKL}E_{KL} \cong C^r_{IJKL}\frac{\partial u_L}{\partial X_K}, \tag{31}$$

where C^r_{IJKL} is defined in equation (30). This equation is analogous to that for the hyperelastic body (equation (16)). The natural elastic moduli $\mathbf{c}^o$ are again changed by the initial deformation which is divided into two parts, the elastic part and the plastic part. The effect of the former $\partial \mathbf{u}^e/\partial \mathbf{X}$, is the same as that for the hyperelastic body. This is seen by comparing the right-hand side of equation (30) (minus the $\mathbf{c}^p$ term) with that of equation (19). The effect of the latter, $\partial \mathbf{u}^p/\partial \mathbf{X}$, is represented by the first term $\mathbf{c}^p$ in equation (30).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The equation of motion (equation (8)) is valid for a body with elastic or plastic predeformation. Substituting the modified constitutive equation (equation (31)), instead of equation (16), into equation (8), we obtain the equation of motion for the incremental displacement **u** of an elastoplastically deformed body

$$\frac{\partial}{\partial X_J}\left[(\delta_{IK}\sigma^i_{JL}+C^r_{IJKL})\frac{\partial u_K}{\partial X_L}\right]=\rho^i\frac{\partial^2 u_I}{\partial t^2}, \tag{32}$$

where C^r_{IJKL} is defined in equation (30). This completes the theory.

IV Plane Waves in Orthotropic Media

For body with homogeneous predeformation, equation (32) is simplified to

$$A^i_{IJKL}\frac{\partial^2 u_K}{\partial X_J \partial X_L}=\rho^i\frac{\partial^2 u_I}{\partial t^2}, \tag{33}$$

where

$$A^i_{IJKL}=\delta_{IK}\sigma^i_{JL}+C^r_{IJKL}. \tag{34}$$

This equation forms the basis for investigating acoustobirefringence in an elastically or plastically predeformed body.

A plane wave propagating in a body in initial coordinates may be represented by

$$u_I=U_I\exp[iK(N_JX_J-Vt)], \tag{35}$$

where U_I is the amplitude, K the wave number, N_J the wave normal, and V the phase velocity of the plane wave. The V and U_K must satisfy the following system of equations,

$$(A^i_{IJKL}N_JN_L-\rho^iV^2\delta_{IK})U_K=0, \tag{36}$$

where V is determined from the characteristic equation,

$$|D_{IK}-\rho^iV^2\delta_{IK}|=0, \tag{37}$$

where

$$D_{IK}=A^i_{IJKL}N_JN_L=(\delta_{IK}\sigma^i_{JL}+C^r_{IJKL})N_JN_L. \tag{38}$$

The **D(N)** is called the acoustical tensor which is a function of the plane wave normal **N**.

For an orthotropic medium, the planes of orthotropic symmetry are taken as the coordinate planes. If the state of stresses is such that the principal stresses σ^i_1, σ^i_2, σ^i_3 are aligned with the coordinate axes, the acoustic tensors are

$$\begin{aligned}
D_{11}&=(C^r_{11}+\sigma^i_1)N_1^2+(C^r_{66}+\sigma^i_2)N_2^2+(C^r_{55}+\sigma^i_3)N_3^2,\\
D_{22}&=(C^r_{66}+\sigma^i_1)N_1^2+(C^r_{22}+\sigma^i_2)N_2^2+(C^r_{44}+\sigma^i_3)N_3^2,\\
D_{33}&=(C^r_{55}+\sigma^i_1)N_1^2+(C^r_{44}+\sigma^i_2)N_2^2+(C^r_{33}+\sigma^i_3)N_3^2,\\
D_{12}&=D_{21}=(C^r_{12}+C^r_{66})N_1N_2,\\
D_{23}&=D_{32}=(C^r_{44}+C^r_{23})N_2N_3,\\
D_{31}&=D_{13}=(C^r_{13}+C^r_{55})N_3N_1,
\end{aligned} \tag{39}$$

where the modified incremental elastic coefficients C^r_{IJKL} are written in Voigt's contracted notation, C^r_{MN}, in which 11 is replaced by 1, 22 by 2, 33 by 3, 23 by 4, 31 by 5, and 12 by 6. The wave speeds V are given by the three roots of equation (37).

V Acoustoelastic Birefringences

From the evaluations of eigenvalues and eigenvectors in equation (37) and (36), we can determine the speeds of waves in a prestressed orthotropic medium propagating in any direction. A few cases of propagational directions which are of interest in applications are discussed as follows.

A Plane Waves in Pure Modes. For plane waves propagating in the direction of the axis of orthotropic symmetry, the motion may be either a pure longitudinal wave or a pure transverse wave. The nine wave speeds V_{IJ} propagating along X_I - axis with polarization in the direction of X_J - axis are listed in equation (40) in matrix form,

$$[\rho^iV^2_{IJ}]=\begin{bmatrix} C^r_{11}+\sigma^i_1 & C^r_{66}+\sigma^i_1 & C^r_{55}+\sigma^i_1\\ C^r_{66}+\sigma^i_2 & C^r_{22}+\sigma^i_2 & C^r_{44}+\sigma^i_2\\ C^r_{55}+\sigma^i_3 & C^r_{44}+\sigma^i_3 & C^r_{33}+\sigma^i_3\end{bmatrix}. \tag{40}$$

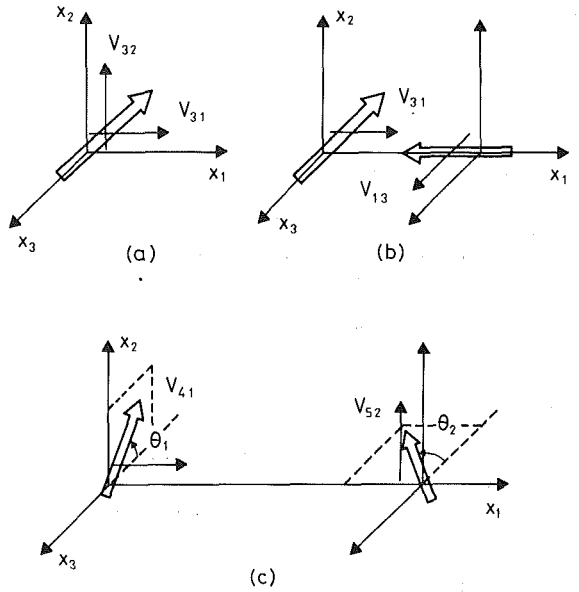


Fig. 3 Birefringence of wave normal (double line arrow) and polarizations (single line arrow) of shear waves in pure modes

Based on the results, two types of birefringent formulae may be derived.

(1) Two shear waves propagate in the same direction, say X_3-axis (Fig. 3(*a*)),

$$\rho^i(V^2_{31}-V^2_{32})=C^r_{55}-C^r_{44}. \tag{41}$$

This is the basis for conventional measurement of initial stresses from acoustoelastic birefringences (Pao et al., 1984).

(2) Two shear waves propagate at perpendicular directions, say X_1-and X_3-axes (Fig. 3(*b*)),

$$\rho^i(V^2_{31}-V^2_{13})=\sigma^i_3-\sigma^i_1. \tag{42}$$

In the absence of initial stresses, $V_{13}=V_{31}=(c^o_{44}/\rho^o)^{1/2}$ for orthotropic materials. The difference in these two wave speeds then measures the initial stresses in the X_1–X_3 plane. This was noted first by Biot (1939), and was proposed recently by Thompson et al. (1986) for measuring residual stresses.

B Obliquely Incident Shear Waves. If the wave normal lies in one of the planes of orthotropic symmetry, one of the two transverse waves is polarized in the direction perpendicular to the symmetrical plane, and hence a pure mode of propagation. Let the wave normal be **N**(0, sin θ_1, cos θ_1). The pure mode shear wave speed is given by

$$\rho^iV^2_{41}(\theta_1)=(C^r_{66}+\sigma^i_2)\sin^2\theta_1+(C^r_{55}+\sigma^i_3)\cos^2\theta_1, \tag{43}$$

where the subscript 4 indicates that the wave normal is in the X_2–X_3 plane (5 indicates X_1–X_3 plane and 6 indicates X_1–X_2 plane), and subscript 1 indicates the direction of displacement. Similarly, when **N**(sin θ_2, 0, $-\cos\theta_2$) is in the X_1–X_3 plane, the speed is (Fig. 3(*c*))

$$\rho^iV^2_{52}(\theta_2)=(C^r_{66}+\sigma^i_1)\sin^2\theta_2+(C^r_{44}+\sigma^i_3)\cos^2\theta_2. \tag{44}$$

where θ_1 and θ_2 are acute angles between the propagational normal and the X_3-axis.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

As first proposed by King and Fortunko (1983), equations (43) and (44) can be used to measure the in-plane stresses.

Subtracting equation (44) from equation (43), and setting $\theta_1 = \theta_2 = \theta$, we find

$$\rho^i[V_{41}^2(\theta) - V_{52}^2(\theta)] = (\sigma_2^i - \sigma_1^i)\sin^2\theta + (C_{55}^r - C_{44}^r)\cos^2\theta. \quad (45)$$

By selecting two proper values for θ, we obtain two simultaneous equations for two incident planes $(\sigma_2^i - \sigma_1^i)$ and $(C_{55}^r - C_{44}^r)$. If the speeds V_{41} and V_{52} at these two angles can be measured accurately, we are able to determine the in-plane stress $(\sigma_2^i - \sigma_1^i)$. This was first proposed by Man and Lu (1987).

C Normal and Oblique Incidence. It should be noted that the variable $(C_{55}^r - C_{44}^r)$ is related to $\rho^i(V_{31}^2 - V_{32}^2)$ by equation (41), and the birefringence of the latter can be measured rather accurately. Therefore, by performing measurement of $V_{41}(\theta)$ and $V_{52}(\theta)$ for only one angle, we obtain another formula to determine initial stress $(\sigma_2^i - \sigma_1^i)$ by combining equation (41) and equation (45)

$$(\sigma_2^i - \sigma_1^i)\sin^2\theta = \rho^i[V_{41}^2(\theta) - V_{52}^2(\theta)] - \rho^i(V_{31}^2 - V_{32}^2)\cos^2\theta, \quad (46)$$

when $\theta = \pi/4$, we have

$$(\sigma_2^i - \sigma_1^i) = 2\rho^i[V_{41}^2(45^o) - V_{52}^2(45^o)] - \rho^i(V_{31}^2 - V_{32}^2). \quad (47)$$

D Slightly Orthotropic Media. For material exhibit slight orthotropy at the natural state, the anisotropic modulus $\mathbf{c}^o$ is split into an isotropic part and an orthotropic part according to equation (20). In contracted indices, the isotropic parts are

$$\begin{aligned}
&c'_{11} = c'_{22} = c'_{33} = \lambda + 2\mu \quad c'_{44} = c'_{55} = c'_{66} = \mu \\
&c'_{123} = \nu_1, \quad c'_{456} = \nu_3, \quad c'_{144} = c'_{255} = c'_{366} = \nu_2, \\
&c'_{112} = c'_{223} = c'_{133} = c'_{113} = c'_{122} = c'_{233} = \nu_1 + 2\nu_2 \\
&c'_{155} = c'_{244} = c'_{344} = c'_{166} = c'_{266} = c'_{355} = \nu_2 + 2\nu_3 \\
&c'_{111} = c'_{222} = c'_{333} = \nu_1 + 6\nu_2 + 8\nu_3.
\end{aligned} \quad (48)$$

We retain the anisotropic second-order constants c''_{IJ} but neglect the anisotropic part of the third-order constants c'_{IJK}. The texture constant p^o_{IJKL} is split into two parts analogously and the isotropic part is

$$p^o_{IJKL} = p_1\delta_{IJ}\delta_{KL} + p_2(\delta_{IK}\delta_{JL} + \delta_{IL}\delta_{JK}). \quad (49)$$

The anisotropic part is also neglected here. With these approximations, we find

$$\begin{aligned}
\rho^i V_{33}^2 &= \sigma_3^i + c''_{33} + c^p_{33} + c'_{33}(1 - e^e) + 4c'_{33}e^e_3 \\
&\quad + c'_{331}e^e_1 + c'_{332}e^e_2 + c'_{333}e^e_3, \\
\rho^i V_{32}^2 &= \sigma_3^i + c''_{44} + c^p_{44} + c'_{44}(1 - e^e) + 2c'_{44}(e^e_2 + e^e_3) \\
&\quad + c'_{441}e^e_1 + c'_{442}e^e_2 + c'_{443}e^e_3, \\
\rho^i V_{31}^2 &= \sigma_3^i + c''_{55} + c^p_{55} + c'_{55}(1 - e^e) + 2c'_{55}(e^e_1 + e^e_3) \\
&\quad + c'_{551}e^e_1 + c'_{552}e^e_2 + c'_{553}e^e_3,
\end{aligned} \quad (50)$$

where $e^e = e^e_1 + e^e_2 + e^e_3$ and

$$\begin{aligned}
c^p_{33} &= 4(p_1 + 2p_2)e^p_3, \\
c^p_{44} &= 2p_2(e^p_2 + e^p_3), \\
c^p_{55} &= 2p_2(e^p_1 + e^p_3).
\end{aligned} \quad (51)$$

Let $V_{T0}^o \equiv (\mu/\rho^o)^{1/2} \cong (V_{31} + V_{32})/2$, we find from equation (41),

$$\begin{aligned}
\frac{V_{31} - V_{32}}{V_{T0}} &= \frac{1}{2\rho^i V_{T0}^2}(C_{55}^r - C_{44}^r), \\
&= \frac{1}{2\mu}(c''_{55} - c''_{44}) + \frac{p_2}{\mu}(e^p_1 - e^p_2) + \frac{1}{2\mu}\left(1 + \frac{\nu_3}{\mu}\right)(\sigma_1^i - \sigma_2^i).
\end{aligned} \quad (52)$$

This is the modified acoustoelastic birefringence formula for a prestressed body with plastic strains. The "texture effect" of polycrystalline materials is contained in the first two terms, the first being the inherent anisotropy of the material at its natural state, and the second being the anisotropy induced by additional plastic deformation. The last term is the stress-induced anisotropy.

VI Summary and Discussions

In this part of the paper, a modified theory of acoustoelasticity for an elastoplastically deformed body is proposed. The plastic strain is assumed small in the formulation. We have adopted the existing theory of small deformations superimposed on a finitely deformed body (equation (8)), and proposed a constitutive equation for the incremental stress in the predeformed body (equation (31)). Based on the modified theory, we have obtained several birefringence formulae which relate the changes of shear wave speeds to plastic strains and initial stresses in the body (equations (41)–(52).

In the development of the constitutive equations (equations (30) and (31), we assume that there exist a natural state (undeformed state) and a relaxed state for the body. The relaxed state is defined as the configuration free of all residual stresses. The predeformed body containing either applied stresses or residual stresses is said to be at its initial state. The relaxed state of a predeformed body may be obtained by sectioning or cutting the specimen around a material point to relieve all stresses at the point, as is done in conventional destructive measurement of residual stresses. The residual stresses in many applications are actually the initial stresses relative to the relaxed state.

We have assumed that the second-order moduli $\mathbf{c}^o$ of the body at the natural state are changed to $\mathbf{c}^r$ at the relaxed state (equations (24), (25)), even both states are free of stresses. This is based on previous observations that acoustoelastic birefringence were changed even after the total removal of external loading in uniaxial tension-compression tests with homogeneous plastic strains. These changes are attributed mainly to the reorientations of polycrystalline grains in a metal to a preferred direction during small plastic deformation. The slight change of the preferred orientations is assumed to be a linear function of plastic strains and is characterized by a set of new material constants called texture constants $\mathbf{p}^o$ (equation (26)). As a further extension, we may assume that in equation (25), $\mathbf{c}^r = \mathbf{c}^o + \mathbf{c}^p$, where $\mathbf{c}^p$ is a general function of plastic strains and other variables.

Additional experimental observations of the change of $\mathbf{c}^o$ to $\mathbf{c}^r$ by homogeneous plastic strains for carbon steel specimens (C1018) are reported in Part II of this paper. The cold rolled steel specimen as received in the laboratory (at the natural state) is slightly orthotropic which we called the inherent anisotropy. From the acoustoelastic birefringent experiments, we are able to measure the texture constant p_2 (equation (49)) in uniaxial tests, and verify the derived results by bending tests. Throughout this investigation, we assume that the third-order moduli of the material at the natural state are not affected by plastic deformation, although the changes, if observable, can be easily incorporated in the theory.

From the relaxed state to the initial state with residual stresses, which are always accompanied with plastic strains, we assume that the moduli $\mathbf{c}^r$ are changed to $\mathbf{C}^r$ (equation (30)) by the elastic strains during unloading (from the initial to the relaxed state). This change is the same as that for a hyperelastic deformation where the moduli $\mathbf{c}^o$ at the natural state are changed to $\mathbf{C}^o$ at the initial state (equation (19)). For the latter case, the change from $\mathbf{c}^o$ to $\mathbf{C}^o$ is rigorously derived on the basis of theory of hyperelasticity. Therefore, the change of $\mathbf{c}^r$ to $\mathbf{C}^r$ in the former can be justified if the deformation from the relaxed state to the initial state with residual stresses is also hyperelastic.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

In the absence of plastic strains, $\mathbf{c}^r = \mathbf{c}^o$ and the incremental elastic coefficient reduced from equation (30) to equation (19).

Finally, we note that when the natural state of the body is unknown or unattainable, and one is interested only to know the residual stresses in the material specimen as received, we may postulate the constitutive relation equation (31) for the incremental stresses of the body at the initial state. A direct substitution of the postulated relation into the equation of motion (equation (8)), which is exact for small incremental disturbance, gives rise to equation (32). The last name equation, containing both initial stresses σ^i and incremental moduli $\mathbf{C}^r$ as unspecified quantities, can then be applied to investigate all problems involving residual stresses or applied stresses. This is in essence the theory proposed by Man and Lu (1987) who started with the equation of motion for the first Piola-Kirchhoff stresses at the final state and treated the initial stress σ^i as a constitutive variable. A detailed comparison of the theory in this paper with previous ones (King and Fortunko, 1983; Biot, 1939; Thompson, 1986; Man and Lu, 1987) is given by Wu (1987).

References

Benson, R. W., and Raelson, V. J., "Acoustoelasticity," *Product Engineering*, Vol. 30, pp. 56–59.

Biot, M. A., 1939, "Nonlinear Theory of Elasticity and the Linearized Case for a Body Under Initial Stress," *Philosophical Magazine*, Vol. 27, Series 7, pp. 468–489.

Crecraft, D. I., 1967, "The Measurement of Applied and Residual Stresses in Metals Using Ultrasonic Waves," *J. of Sound Vibr.*, Vol. 5, pp. 173–192.

Hirao, M., and Pao, Y. H., 1985, "Dependence of Acoustoelastic Birefringence on Plastic Strains in a Beam," *J. Acoust. Soc. Am.*, Vol. 77, pp. 1659–1664.

Iwashimizu, Y., and Kubomura, K., 1973, "Stress-Induced Rotation of Polarization Directions of Elastic Waves in Slightly Anisotropic Materials," *Int. J. Solids Structures*, Vol. 9, p. 99.

Johnson, G. C., 1981a, "Acoustoelastic Theory for Elastic-Plastic Materials," *J. Acoust. Soc. Am.*, Vol. 70, pp. 591–595.

Johnson, G. C., 1981b, "On the Applicability of Acoustoelasticity for Residual Stress Determination," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 791–795.

Johnson, G. C., 1983, "The Effect of Plastic Deformation on the Acoustoelastic Response of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 689–691.

King, R. B., and Fortunko, C. M., 1983, "Determination of In-Plane Residual Stress States in Plates Using Horizontally Polarized Shear Waves," *J. Appl. Phys.*, Vol. 54, p. 3207.

Kobayashi, M., 1987, "Theoretical Study of Acoustoelastic Effects Caused by Plastic Anisotropy Growth," *Int. J. Plasticity*, Vol. 3, pp. 1–20.

Man, C. S., and Lu, W. Y., 1987, "Towards an Acoustoelastic Theory for Measurement of Residual Stress," *J. of Elasticity*, Vol. 17, pp. 159–182.

Pao, Y. H., and Gamer, U., 1983, "A Theory of Acoustoelasticity for Elasto-Plastic Solids," Material Science Center report no. 5121, Cornell University, Ithaca, New York.

Pao, Y. H., Sachse, W., and Fukuoka, H., 1984, "Acoustoelasticity and Ultrasonic Measurements of Residual Stresses," *Physical Acoustics*, Vol. 17, W. P. Mason and R. N. Thurston, eds., Academic Press, New York, pp. 62–143.

Pao, Y. H., and Gamer, U., 1985, "Acoustoelastic Waves in Orthotropic Media," *J. Acoust. Soc. Am.*, Vol. 77, No. 3, pp. 806–812.

Pao, Y. H., 1985, "Theory of Acoustoelasticity and Acoustoplasticity," *Solid Mechanics Research for Quantitative Non-Destructive Evaluation*, J. D. Achenbach and Y. Rajapakse, eds., Martinus Nijhoff, Dordrecht, p. 257.

Thompson, R. B., 1986, "Angular Dependence of Ultrasonic Wave Propagation in a Stressed, Orthotropic Continuum: Theory and Application to the Measurement of Stress and Texture," *J. Acoust. Soc. Am.*, Vol. 80, No. 3, pp. 921–931.

Thurston, R. N., and Brugger, K., 1964, "Third-Order Elastic Constants and the Velocity of Small Amplitude Elastic Waves in Homogeneously Stressed Media," *Phys. Rev.*, Vol. A133, pp. 1604–1610.

Toda, H., Fukuoka, H., and Ohmori, H., 1982, "Variations in Texture-Induced Acoustic Anisotropy due to Compressive Deformations (Influence of Strain-Rate and Temperature)," *J. Soc. Mater. Sci. Jpn.*, Vol. 31, p. 36.

Tokuoka, T., and Iwashimizu, Y., 1968, "Acoustical Birefringence of Ultrasonic Waves in Deformed Istropic Elastic Materials," *Int. J. Solids and Structures*, Vol. 4, pp. 383–389.

Tokuoka, T., and Saito, M., 1969, "Elastic Wave Propagations and Acoustical Birefringence in Stressed Crystals," *J. Acoust. Soc. Am.*, Vol. 45, pp. 1241–1246.

Toupin, R. A., and Bernstein, B., 1961, "Sound Waves in Deformed Perfectly Elastic Materials. Acoustoelastic Effect," *J. Acoust. Soc. Am.*, Vol. 33, pp. 216–225.

Wong, T. Eric, and Johnson, George C., 1985, "The Acoustoelastic Response of an Aluminum Plate During Elasto-Plastic Deformation," *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, eds., Vol. 5B, Plenum, New York, p. 1397.

Wu, T. T., 1987, "Theory of Acousto-Plasticity and Ultrasonic Measurements of Residual Stresses," Ph.D. Thesis, Cornell University, Ithaca, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Tsung-Tsong Wu[2]

Masahiko Hirao[3]

Yih-Hsing Pao
Fellow ASME.

Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853

Acoustoelastic Birefringences in Plastically Deformed Solids: Part II—Experiment[1]

Ultrasonic experiments are preformed on a carbon steel specimen to determine the change of elastic wave speeds by plastic strains and residual stresses in the specimen. Under repeated uniaxial loadings, the acoustoelastic birefringence was found to be a linear function of plastic strains at various states of total unloading. In elastoplastic bendings, the acoustoelastic birefringes are dependent on the natural anisotropy, plastic strains, and residual stresses as predicted by the theory of Part I. The uniaxial residual stress and plastic strain in the beam are determined by acoustoelastic experiments.

In the second part of the paper, we report a series of experiments which were designated to investigate the dependence of acoustoelastic birefringences on initial stresses (applied stress and residual stress) as well as plastic strains. Two types of loadings were applied to carbon steel specimens, the uniaxial tensile-compressional loading and the four-point bending. The first type of loading is supposed to generate homogeneous plastic deformation, at least in the central portion of the specimen. The bending test develops nonhomogeneous plastic strains in the beam.

The specimens were cut from large, cold rolled plates of carbon steel (C1018). The length of one group of specimens, designed as group A, is parallel to the rolling direction of the plate; the length of the other group, designated as group B, is perpendicular to the rolling direction. These two groups of specimens exhibited different anisotropic properties as received in the laboratory (at the natural state) and responded differently to ultrasonic excitation in plastic deformation.

Ultrasonic waves were transmitted through the thickness of the specimen at various stages of the loading-unloading process (the initial states), and the shear wave speeds at two perpendicular polarizations were measured by the pulse-echo-overlap technique. The difference of these two shear wave speeds, known as the acoustoelastic birefringence, or simply acoustobirefringence, was recorded at each initial state.

For the case of bending with inhomogeneous plastic strains, the beam specimen was also sectioned, after total unloading, to relieve the residual stress. Ultrasonic tests were performed on the specimen at both initial state and relaxed state. Strain gauge readings were also taken before and after the sectioning so as to measure the residual stresses as is done by the conventional "destructive" method.

In the next section, we discuss the effect of texture on acoustoelastic birefringence from the point of view of microbehavior of materials, and summarize the formulae needed for experiments. Section II then presents the results for uniaxially loaded specimens, and Section III for a beam in pure bending. The last section concludes both parts, Part I and Part II, of this paper.

I Acoustoelastic Birefringence and Plastic Deformation

A Texture Effect. A polycrystalline specimen is expected to be isotropic if its grains are randomly oriented. From a macroscopic point of view, the anisotropy of the crystals vanishes on the average of many grains. A truly random arrangement of the crystals seldom exists in a commercially available metal specimen, however, because manufacturing processes tend to align the grains toward one or two preferred orientations. The result is known as the texture or the preferred orientation of a metal (Reed-Hill, 1964).

Because most of the polycrystalline metals have a preferred orientation, they tend to be slightly anisotropic, and the degree of anisotropy depends on the degree of plastic working during the manufacturing or forming processes. For example, a cold rolled metal plate usually exhibits a preferred orientation along the mechanical rolling direction of the plate, and a wire has a preferred orientation along the drawing axis. The geometry of the forming process seems to play an important role in the formation of the texture in metals.

Acoustoelastic experiments with metals in elastic deformation showed that the preferred orientations remained un-

[1]Report No. 6292, Material Science Center, Cornell University.

[2]Presently at the Institute of Applied Mechanics, National Taiwan University, Taipei, Taiwan, R.O.C.

[3]On leave from the Faculty of Engineering Science, Osaka University, Osaka, Japan.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, Dallas, Texas, Nov. 25–30, 1990.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 30, 1988; final revision, Nov. 14, 1989.

Paper No. 90-WA/APM-42.

Copyright © 1991 by ASME

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

changed during elastic deformation. On the other hand, it was found that the preferred orientation was perturbed by a small amount of plastic deformation (less than one percent) in some metals (Johnson, 1981; Toda et al., 1982; Hirao and Pao, 1985; Wong and Johnson, 1985). Since residual stresses result from inhomogeneous plastic deformations in a material with texture, it is necessary to study first the interaction between the plastic deformation and the preferred orientation.

Experimental investigations of the effect of plastic deformation on the acoustoelastic response of some typical materials have been reported by several investigators. Johnson (1981) investigated the effects of plastic deformation on the longitudinal wave velocities for three different aluminum alloys. Toda et al. (1982) studied the strain rate and temperature effects on the texture-induced acoustic anisotropy of uniaxial deformation. Their results showed that for mild steel (C : 0.17 percent) and pure aluminum, the acoustobirefringences are linearly proportional to plastic strains as large as 10 percent in magnitude. But for pure copper, they are highly nonlinear. Hirao and Pao (1985) conducted uniaxial loading and bending experiments with aluminum specimens (Al 6061-T6). Their results showed that the acoustoelastic birefringences are nonlinearly dependent on the plastic strains. For plastic strains less than 0.5 percent, however, the dependence may be approximated by two linear relations, one for tension and the other for compression. Wong and Johnson (1985), performed experiments on a different kind of aluminum rolled plate (5086-H32) in uniaxial tests. Their results showed that the change of the wave velocities at the unstressed state (relaxed state) are linearly dependent on the plastic strain.

A statistical theory for the change of preferred orientations was proposed recently by Johnson (1985) and Sayer (1987). In the theory, however, the changes are not related to plastic strains.

B Modified Birefringent Formula of Acoustoelasticity. To incorporate the effect of texture, a theory of acoustoelasticity is presented in Part I. The theory is embodied in the general equation of motion for the superposed dynamic disturbance on a predeformed body, equation (8), and the postulated constitutive equations, equations (30) and (31), for the incremental stresses due to dynamic disturbance. The texture effect of the metal specimen as received in the laboratory is included in the anisotropic moduli, $\mathbf{c}^o$, at the natural state. The change of texture due to additional plastic strains in the laboratory is represented by the perturbation $\mathbf{c}^p$ in equation (25) and the moduli are changed to $\mathbf{c}^r$ (at the relaxed state). The addition of initial stresses change $\mathbf{c}^r$ to $\mathbf{C}^r$.

For plane waves propagating in the X_3-direction (the thickness direction of a specimen), which coincides with one of the axes of symmetry of a slightly orthotropic material, the acoustoelastic birefringence is shown to be a linear function of initial stresses and the plastic strains (equation (52) of Part I),

$$\frac{V_{31}-V_{32}}{V_{T0}}=\frac{\Delta V}{V_{T0}}=\alpha_o+C_t(e_1^p-e_2^p)+C_a(\sigma_1^i-\sigma_2^i), \quad (1)$$

where

$$\alpha_o=\frac{1}{2\mu}(c_{55}''-c_{44}''),$$
$$C_t=p_2/\mu, \quad (2)$$
$$C_a=\frac{1}{2\mu}\left(1+\frac{\nu_3}{\mu}\right).$$

The first term, α_o, in equation (1) is the birefringence of the original anisotropy measured in the natural state; the second term represents the change of texture as influenced by plastic strains; the last term is the acoustoelastic effect which depends on the initial stresses. We call C_a the acoustoelastic coefficient, and C_t the acoustoplastic coefficient.

In the case of uniaxial testing, the principal stress σ_1^i and principal plastic strain e_1^p are measured along the longitudinal axis (axis of loading) of the specimen, and $\sigma_2^i=0$. For slightly orthotropic materials, it is reasonable to assume that e_2^p is approximately equal to $-e_1^p/2$. Thus, the birefringent formula reduces to

$$\frac{\Delta V}{V_{T0}}=\alpha_o+\frac{3}{2}C_te_1^p+C_a\sigma_1^i. \quad (3)$$

From equation (3), it is clear that the original birefringence α_o, is measured in the natural state

$$\left(\frac{\Delta V}{V_{T0}}\right)^o=\alpha_o=\frac{1}{2\mu}(c_{55}''-c_{44}''). \quad (4)$$

The acoustoelastic coefficient C_a can then be calibrated by measuring the response to the elastic deformation ($e_1^p=0$)

$$\left(\frac{\Delta V}{V_{T0}}\right)^i=\alpha_o+C_a\sigma_1^i. \quad (5)$$

Finally, the acoustoplastic coefficient C_t is determined from data in the relaxed state ($\sigma_1^i=0$)

$$\left(\frac{\Delta V}{V_{T0}}\right)^r=\alpha_o+\frac{3}{2}C_te_1^p. \quad (6)$$

In the case of elastoplastic bending without shearing force, the dominant stress in the beam is the normal stress in the length direction of the beam. Equations (3)–(5) are also applicable if the X_1-axis is in the length direction.

II Acoustoelastic Birefringence in Uniaxial Loading and Unloading

To investigate the influences of texture in homogeneous plastic deformation, standard uniaxial tension-compression tests are performed with steel specimens. Before conducting the tests, two cross-ply strain gages are mounted on each specimen to monitor the longitudinal (uniaxial loading direction) and transverse strains. A homemade dual-polarization shear wave transducer is also mounted to measure the changes of the birefringence (the difference of two shear wave speeds, equation (1)).

After being cut from a large steel plate, the specimen is assumed to be free of stress and strain and is in its natural state. It is then loaded either in tension or compression in step loadings. At each step, the loading machine is stopped for taking the birefringence data of the specimen through the echo-overlap system (Hsu, 1974) with the accuracy about one part in 10^5. Details of the ultrasonic measurements were described previously (Hirao and Pao, 1985). The testing equipment and data acquisition system is shown schematically in Fig. 1.

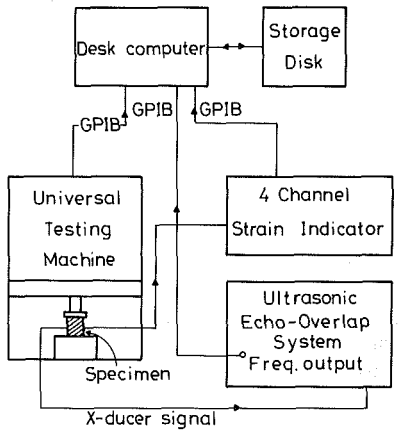


Fig. 1 Block diagram of the data acquisition system

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

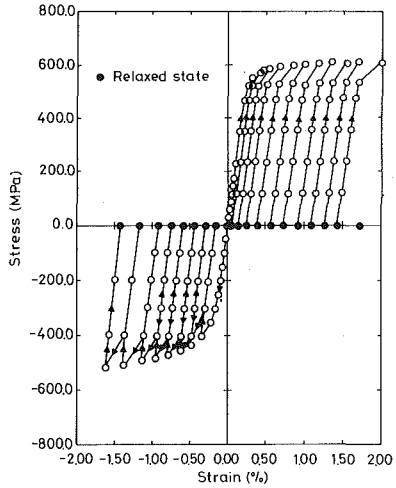


Fig. 2 Stress strain curve for the A-group steel specimens (length direction parallel to the rolling direction) in tension (upper half) and compression (lower half)

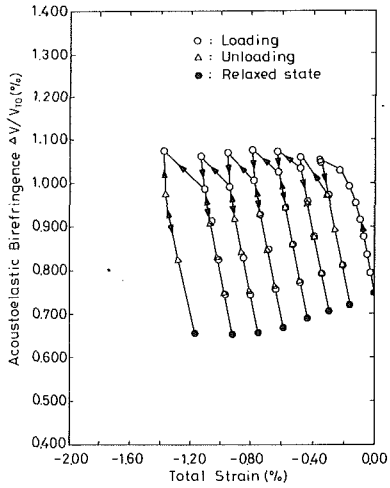


Fig. 3 Acoustoelastic birefringence history of A-group specimen (sample 1) during *compressive* elastoplastic cyclic loading

While the loading is in the elastic range, the anelastic effect is negligible, but the taking of birefringence data must be delayed for five minutes or longer to ensure the mechanical equilibrium between the specimen and the load cell of the testing machine (Shimadzu AG-10TA). To monitor the amount of plastic strain for each cycle of loading and unloading, the load-displacement curve is displayed on the terminal of a desktop computer (HP9836). From the load-displacement curve, it is easier to control the amount of plastic strain in each cycle.

During loading and unloading, the contact pressure between the specimen and the transducer is changed by the lateral deformation. Hence, to maintain a constant contact pressure, the amplitude of a particular ultrasonic pulse echo is maintained constant by adjusting screws in the holding fixture of the transducer and the specimen.

A Acoustoelastic Response in Repeated Elastoplastic Deformation. Four specimens of group A (length parallel to the rolling direction) are used in the uniaxial tests, two for tension and two for compression. The dimensions for all specimens are 12.5 mm in thickness, 16.0 mm in width, and 25.4 mm (compression) in length. Shown in Fig. 2 are the typical history of the stress-strain relation for a specimen under repeated loading-unloading. In the case of compressive loading, for example, the specimen was loaded, from the origin O (natural state), at some steps (open circles), passing the elastic limit Y, to the plastic region, and then unloaded to the relaxed state (solid circles). All the other cycles of loading were then started from their respective relaxed states (solid circles). At each loading or unloading step, acoustobirefringence experiments were conducted.

From the elastoplastic stress-strain relation (Fig. 2), it was found that the yield behavior of the A-group specimens for tension is different from that for compression, the yield limit being higher in tension than that in compression. The work hardening in compressive strain, however, is larger than that in tensile strain.

Figure 3 shows the relations between the birefringence and the total strain of one of the A-group compressive specimens (loading direction parallel to the rolling direction). Because of the original anisotropy of the cold rolled plate, the original birefringences (α_o) of both specimens do not vanish (about 0.7 percent). As the compressive force increases, the birefringences (open circles) also increase (going upward) because of the negative value of the acoustoelastic constant C_a. In Fig. 3, the open circles represent the birefringences measured during the loading of the specimen, whereas, the open triangles are those measured during unloading. The solid circles are the birefringence measurements made in the relaxed states ($\sigma_1^i = 0$).

As can be seen in Fig. 3, within the elastic range, the birefringences measured during unloading (open triangles) and loading (open circles) are approximately equal, and hence, the birefringence measurements of all the other A-group specimens are measured only during the loading process.

From Fig. 3, we found that the birefringences in the relaxed states (solid circles) are different from that in the natural state (α_o), and are dependent on the amount of the compressive plastic strain. Moreover, at small plastic strain (less than 0.6 percent), the dependency is almost linear. During each cycle of reloading, the relation between the birefringence and the elastic strain of each cycle is linear and the slopes of all reloading lines are about the same. This is consistent with the theory equation (2) and previous observation that the acoustoelastic constant (C_a) is insensitive to the plastic deformation (Hirao and Pao, 1985).

Four specimens of group B (length perpendicular to the rolling direction) were also tested under the same condition, two in tension and two in compression. Only two of these four specimens (one tension and one compression) are tested for the complete history of loading and unloading, the other two are used only to obtain the changes of acoustoelastic birefringences in relaxed states.

B Determination of Acoustoelastic and Texture Constants. The general relationships between the acoustobirefringences and the plastic strains (tension and compression) of carbon steel (C : 0.18 percent) at the relaxed state are shown in Fig. 4 (A group) and Fig. 5 (B group). Solid and open circles are used to distinguish data taken from different specimens, two in tension, two in compression.

For A-group specimens, the results in Fig. 4 show that within the range of small plastic strain (less than 0.6 percent in tension or compression), the relation between the birefringence and the plastic strain is approximately linear. For B-group specimens, the results in Fig. 5 show that the dependence of the birefringence on the tensile plastic strain is different from that on the compressive plastic strain. However, for small plastic strain, a linear relation for tension and another one for compression may be assumed.

Based on equation (6), the acoustoplastic coefficient C_l is given by the slope of the linear relation between the acoustobirefringence, $\Delta V/V_{T0}$, and small plastic strain, e_1^p. By using a linear least square fitting, the acoustoplastic coefficients for both A and B group specimens are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

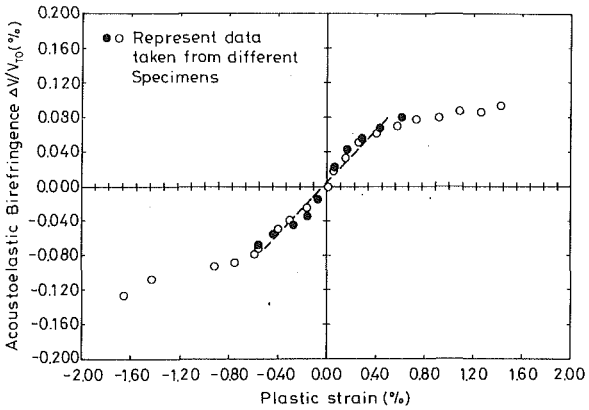


Fig. 4 Acoustoplastic birefringence of A-group specimens in tension and in compression

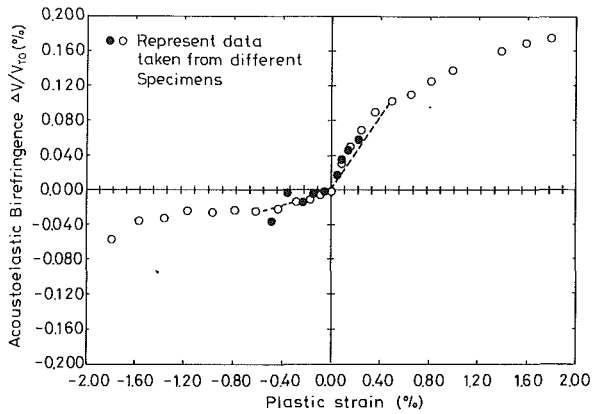


Fig. 5 Acoustoplastic birefringence of B-group specimens in tension and in compression

A group $C_t = 0.104$ (tension and compression)
B group $C_t = 0.141$ (tension)
$C_t = 0.016$ (compression).

The second and third-order elastic constants of group-A specimens were determined by measuring the changes of ultrasonic wave velocities (both P and S-waves) induced by elastic deformation in the specimens. Three specimens of group A were used in this experiment. To determine the Lamè constants λ, μ in the natural state, because the specimens are slightly orthotropic, the average values of the S-wave speeds at the two directions of polarization were used. The measured Lamè constants and the third-order elastic constants in this experiment are

$$\mu = 81.0 \pm 0.2 \text{ GPa}, \quad \lambda = 113.0 \pm 0.2 \text{ GPa}$$
$$\nu_1 = -210 \pm 40 \text{ GPa}$$
$$\nu_2 = -344 \pm 11 \text{GPa}$$
$$\nu_3 = -193 \pm 7 \text{ GPa}.$$

Although the acoustoelastic constant C_a as defined in equation (4.2) can be calculated from the values of second and third-order constants, it is usually determined by measuring the applied stress σ_1^i versus $\Delta V/V_{T0}$ in the elastic range. The Young's moduli and the corresponding C_a are obtained from uniaxial tests as follows:

A group
Tension: $C_a = -8.60 \times 10^{-6} (\text{MPa}^{-1})$(tension)
$E = 217.55$ GPa (tension)
Comp.: $C_a = -8.34 \times 10^{-6} (\text{MPa}^{-1})$(compression)
$E = 216.2$ GPa (compression)

B group
Tension: $C_a = -8.01 \times 10^{-6} (\text{MPa}^{-1})$(tension)
$E = 215.0$ GPa (tension)
Comp.: $C_a = -8.15 \times 10^{-6} (\text{MPa}^{-1})$(compression)
$E = 213.8$ GPa (compression).

The preceding value for C_a of group-A specimen can be compared with that calculated from the second and third-order constants, which are determined as $\mu = 81.0$ GPa, $\nu_3 = -193$ GPa. Hence, we find $C_a = -8.53 \times 10^{-6}$ MPa^{-1} (equation (2)).

It should be noted at this stage that outside the limited range of small plastic strains, the assumption of a linear relation between the acoustobirefringence and plastic strain is not valid.

III Acoustoelastic Birefringence in Elastoplastic Bending

Uniaxial tests on carbon steel specimens (C:0.18 percent) showed that the acoustoelastic birefringences in the relaxed

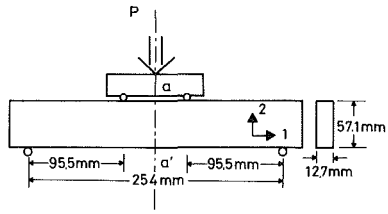


Fig. 6 Geometry of the four-point beam bending test

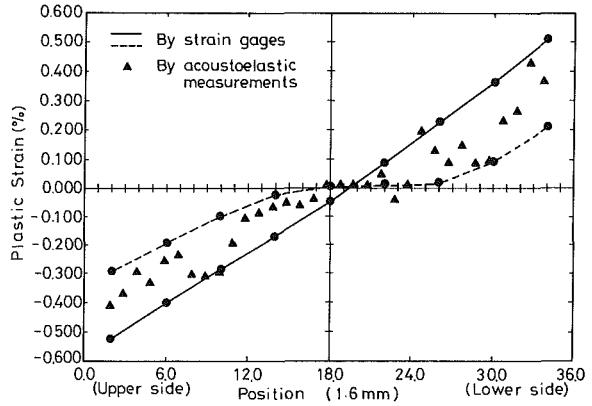


Fig. 7 Distribution of total strains $e = e^e + e^p$ (solid line) at maximum loading and plastic strains e^p (dashed line) along the center line of the bending specimen A

states are changed by homogeneous plastic strains, and the acoustoplastic birefringences are linearly dependent on small plastic strains. Further to test the effect of nonhomogeneous plastic deformations, four-point bending experiments were conducted with the same steel sample. The experimental techniques are the same as those reported in a previous paper (Hirao and Pao, 1985).

A Acoustoelastic Experiments in Bending Tests. Two four-point bending tests were conducted, one for a specimen of group A and one for that of group B. The dimensions of the bending specimens (A and B) are shown in Fig. 6.

To measure the distribution of the total strain (e_1) at the centerline (a-a') of the beam, nine strain gages were mounted with equal spacing (7.1 mm). They also serve to measure residual stresses by sectioning the beam around each gage. The total strains in the beam (specimen A) at the maximum loading and permanent strains after sectioning of the beam are shown in Fig. 7. To avoid the influence of the strain gages on the birefringence measurements (wave reflections) at the initial state (after plastic bending), the nine strain gages were re-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

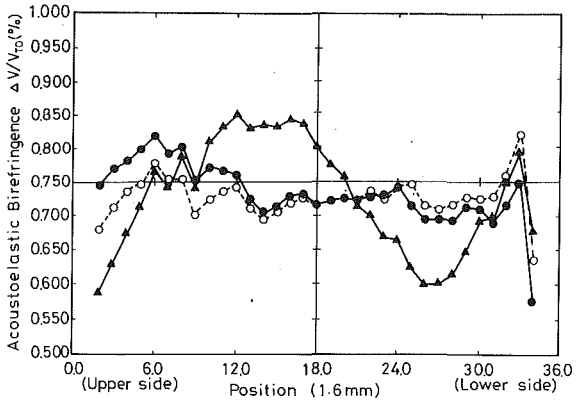


Fig. 8(*a*) Distribution of acoustoelastic birefringences along the center line of the bending specimen A at three different states, natural state (solid circles), initial state (triangles), and relaxed state (open circles)

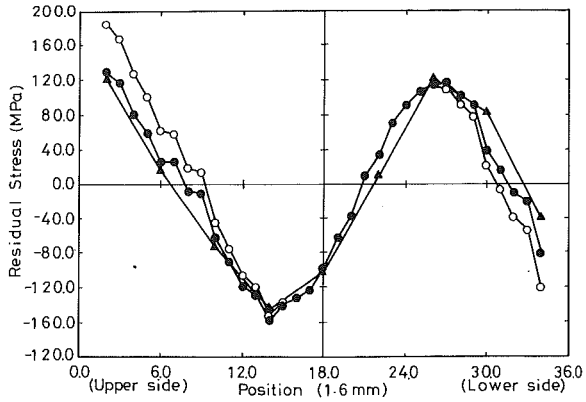


Fig. 9(*a*) Distribution of residual stresses in bending specimen A determined by destructive method (triangles) and by ultrasonic measurements (open circles are based on acoustoelastic theory, solid circles on acoustoplastic theory)

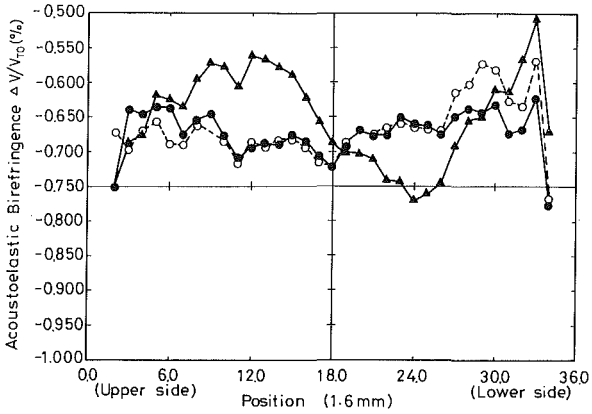


Fig. 8(*b*) Distribution of acoustoelastic birefringences along the center line of the bending specimen B at three different states, natural state (solid circles), initial state (triangles), and relaxed state (open circles)

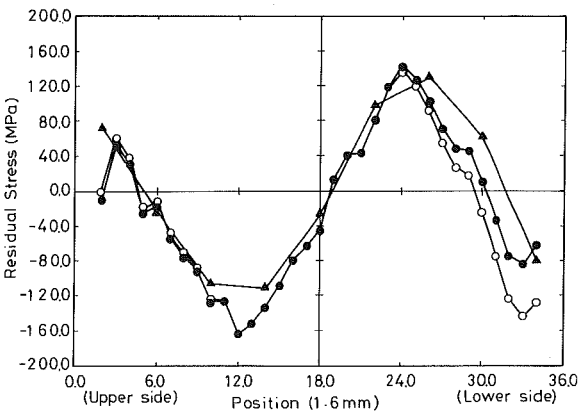


Fig. 9(*b*) Distribution of residual stresses in bending specimen B determined by destructive method (triangles) and by ultrasonic measurements (open circles are based on acoustoelastic theory, solid circles on acoustoplastic theory)

moved, and a new set of strain gages were mounted on again after the measurements. Similarly, before the birefringence measurements at the relaxed state (after cutting), the strain gages were removed.

Along with the four-point bending tests, the acoustobirefringences were measured. The side of the bending specimen without strain gages was polished carefully to achieve accurate measurement of the birefringences. The measurements were taken along the center line (*a-a'* in Fig. 6) at a spacing of 1.6 mm. Hence, there are 34 data points across the depth of the beam from the upper to the lower side.

Before the bending, the beam is in its natural state, and the inherent birefringences of the beam are measured according to equation (4). After the beam is bent plastically, both inhomogeneous plastic strains and residual stresses are induced, and the acoustobirefringences in this initial state, $(\Delta V/V_{T0})^i$, are determined according to equation (3) or equation (5). As the residual stresses at the initial state are removed by sectioning the beam, the acoustobirefringences in this relaxed state, $(\Delta V/V_{T0})^r$, are determined by equation (6).

Figures 8 show the acoustobirefringences measured at the natural states (solid circles), initial states (solid triangles), and relaxed states (open circles) of the beam specimens A and B, respectively. For specimen A (Fig. 8(*a*)), the inherent acoustobirefringences in the natural states (solid circles) fluctuates around 0.75 percent, with considerable scattering due to manufacturing processes. After the bending, the birefringences in the initial state (solid triangles) are changed by both residual stresses and plastic deformations. In the relaxed state (after cutting), the measurements (open circles) showed that at the outer region of the beam, the birefringences are changed by the plastic strains. As compared with those in the natural state, the birefringences in the relaxed state are decreased by the compressive plastic strains in the upper part of the beam and increased by the tensile plastic strains in the lower part of the beam. In the central region where the material points experience only elastic deformation, the birefringences in the relaxed state are very close to those in the natural state.

For specimen B (Fig. 8(*b*)), the inherent birefringences scattered around the value of -0.67 percent. The other results are similar to those of specimen A. The changes of the birefringences at the upper part of the beam (compressive plastic deformation) are small, which is in agreement with the results under uniaxial loading.

B Effect of Plastic Strain on Residual Stress Measurement. Determination of residual stresses in a beam by ultrasonic measurements is based on the birefringent formulae equations (3)–(5). The values for C_t, C_a are determined from the uniaxial experiments and given in the previous section.

Shown in Figs. 9(*a*) and 9(*b*) are the distribution of stresses in the bending specimens A and B, respectively. In these two figures, the solid triangles are the residual stresses measured by the strain gages and known value of Young's modulus. The open circles are the residual stresses calculated from the hyperelastic birefringent formula (equation (5)), and the solid circles are those from the modified birefringent formula (equation (3)).

In both figures, the results based on the formula without

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

plastic strains (equation (5), open circles) deviate substantially from the residual stresses measured by strain gages (solid triangles). The differences between these two measurements are approximately proportional to the amount of the plastic strains locked in the bent beam. The results based on the modified formula (equation (3)) show much better agreement. Without the corrections for plastic strain, the error can be as large as 60 MPa in these tests.

C Measurement of Plastic Strains. Since the acousto-plastic coefficients C_t of the steel specimens were determined from the uniaxial experiments, the distribution of the plastic strains in the bending tests can be obtained by subtracting equation (6) from equation (4), i.e., the differences of the birefringences between the natural state and the relaxed state. Solid triangles in Fig. 7 represent the distribution of plastic strains in bending specimen A as determined by acoustoelastic experiment. Although the data are scattering, the trend and magnitude of the plastic strains are the same as that measured by strain gages. The scattering of the data may be due to the slight unequal cutting depth around the strain gages in the sectioning process.

IV Conclusions

In the second part of the paper, we report two experiments, uniaxial loading and pure bending, to test the theory of acoustoelasticity presented in Part I. The theory is based on a new constitutive equation for the incremental stress (Second Piola-Kirchhoff stress)

$$T_{IJ} = C^r_{IJKL} E_{KL} \cong C^r_{IJKL} \frac{\partial u_L}{\partial X_K}, \tag{I.31}$$

and the balance equation of momentum for the superposition of small disturbance on a finitely deformed body,

$$\frac{\partial}{\partial X_J}\left(T_{IJ} + \frac{\partial u_I}{\partial X_K}\sigma^i_{JK}\right) = \rho^i \frac{\partial^2 u_I}{\partial t^2}. \tag{I.8}$$

Since the latter is founded on the continuum theory without other assumptions except for the small disturbance, the test is on the constitutive equation, especially the formulation of the coefficients $\mathbf{C}^r$ (equations (I.29), (I.30)), which consists of the usual second-order constant $\mathbf{c}^o$, the third-order constant $\mathbf{c}^o_3$, and the texture constants $p^o_{\alpha\beta\gamma\delta}$ (equation (I.24)) at the natural state for polycrystalline materials.

Experimentally, we have shown in this part that all material constants ($\mathbf{c}^o$, $\mathbf{p}^o$, $\mathbf{c}^o_3$) needed for acoustoelastic experiments can be obtained by uniaxial test of a plate specimen with slight orthotropy (rolled steel plate). The second-order constant, $\mathbf{c}^o$, indeed is changed by uniform plastic strains in specimens free of stresses, and $\mathbf{c}^p$ is a linear function of plastic strains as assumed in the theory provided that the initial deformation is small.

Both the original anisotropy (the first term in equation (I.52) or α_o in equation (II.1)) and the induced anisotropy by plastic deformation (the second term in equation (I.52) or (II.1)) are very small (less than one percent) and are neglected in many engineering applications. In the ultrasonic measurement of residual stresses, however, they are of the order of magnitude as that induced by initial stresses (the third term of equation (I.52) or (II.1)). The acoustoelastic experiment on the steel beam in pure bending determines all three parts precisely (Figs. II.8). The experiment also measures quantitatively the plastic strains and texture effect on acoustoelastic birefringence.

As discussed in several previous papers and in Part I, the theory in Part I (Section V) shows that the in-plane values of $\mathbf{C}^r$ and σ^i can be measured by several ultrasonic beams propagating at different directions in a body with initial deformation. Various methods of ultrasonic measurements of residual stresses were compared by Clark et al. (1987). When the experimental techniques are perfected, it is possible to determine the initial stress, σ^i, and the effective moduli, $\mathbf{C}^r$, directly by ultrasonics.

Acknowledgment

The research reported in Part I and Part II was supported in part by the National Science Foundation through a grant to the Materials Science Center, Cornell University. The experimental work reported in Part II was conducted at the Institute of Applied Mechanics, National Taiwan University.

References

Clark, A. V., Moulder, J. C., Mignogna, R. B., and Delsanto, P. P., 1987, "A Comparison of Several Ultrasonic Techniques for Absolute Stress Determination in the Presence of Texture," *Solid Mechanics Research for Quantitative Non-Destructive Evaluation*, J. D. Achenbach and Y. Rajapakse, eds., Martinus Nijhoff Publishers, Dordrecht, pp. 345–360.

Hirao, M., and Pao, Y. H., 1985, "Dependence of Acoustoelastic Birefringence on Plastic Strains in a Beam," *J. Acoust. Soc. Am.*, Vol. 77, pp. 1659–1664.

Hsu, Nelson N., 1974, "Acoustical Birefringence and the Use of Ultrasonic Waves for Experimental Stress Analysis," *Experimental Mechanics*, pp. 169–176.

Johnson, G. C., 1981, "On the Applicability of Acoustoelasticity for Residual Stress Determination," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 791–795.

Johnson, G. C., 1985, "Acoustoelastic Response of a Polycrystalline Aggregate with Orthotropic Texture," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 659–663.

Reed-Hill, R. E., 1964, *Physical Metallurgy Principles*, D. Van Nostrand Company, Princeton, N.J.

Sayer, C. M., 1987, "Ultrasonic Determination of Texture and Residual Stress in Polycrystalline Metals," *Solid Mechanics Research for Quantitative Non-Destructive Evaluation*, J. D. Achenbach and Y. Rajapakse, eds., Martinus Nijhoff Publishers, Dordrecht.

Toda, H., Fukuoka, H., and Ohmori, H., 1982, "Variations in Texture-Induced Acoustic Anisotropy due to Compressive Deformations (Influence of Strain-Rate and Temperature)," *J. Soc. Mater. Sci. Jpn.*, Vol. 31, p. 36.

Wong, T. Eric, and Johnson, George C., 1985, "The Acoustoelastic Response of an Aluminum Plate During Elasto-Plastic Deformation," *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, eds., Vol. 5B, Plenum, New York, p. 1397.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

R. J. Gu
Asst. Professor

J. D. Hovanesian
Professor and Chair

Y. Y. Hung
Professor

Department of Mechanical Engineering,
Oakland University,
Rochester, MI 48309-4401

Calculations of Strains and Internal Displacement Fields Using Computerized Tomography

A novel treatment of computerized tomography is developed to calculate strains and displacements at internal nodes from known boundary displacements. The formulation is blended with finite element numerical scheme. Numerical simulations are performed to verify accuracy of this new technique using several plane problems whose solutions exist in closed form. Excellent agreement shows the potential of developing the tomographic technique into a hybrid numerical-experimental method for solving mechanics problems.

1 Introduction

In contrast to conventional projection radiography, in which a two-dimensional projection of a three-dimensional object is made, computerized tomography reconstructs a cross-sectional view of a selected transverse plane. The tomographic imaging technique allows a quantitative map of the X-ray attenuation coefficient in any transverse plane to be reconstructed. Reconstruction of this two-dimensional map involves intensive numerical computation on a large number of empirical data collected from X-ray detectors.

Applications using similar mathematical basis have been found in diverse fields. Among those are holographic interferometry by Sweeney and Vest (1973), ultrasonic transmission by Anderson and Kak (1982), and nuclear magnetic resonance by Hennig (1988), to name just a few. In this investigation we develop a numerical scheme of computerized tomography for the area of solid mechanics in which strains and internal displacements may be determined from known boundary displacements regardless of material properties. Using finite element, boundary element, or other computational techniques, the determination of displacement and strain requires precise knowledge of material properties besides boundary conditions. Although no electromagnetic waves are utilized here, we formulate the problem assuming that there are imaginary rays penetrating the object from many different angles. The interpolation technique commonly used in isoparametric finite element method is incorporated in the implementation to take care of cases where the object of concern involves irregular geometry. Numerical simulation is demonstrated using several plane problems whose displacement fields can be obtained exactly in closed form.

2 Formulation and Numerical Implementation

Figure 1 schematically depicts the problem of concern, in which a ray penetrates the object at two arbitrary boundary point P and Q having known displacements u and v in the x and y directions, respectively. The imaginary ray may be represented by a line equation:

$$y = mx + b \tag{1}$$

where m is the slope and b is the y-intercept at $x = 0$. Since $du = \partial u/\partial x\, dx + \partial u/\partial y\, dy$, integration of du along the imaginary ray yields

$$u_Q - u_P = \int_{x_P}^{x_Q} \left(\frac{\partial u}{\partial x} + m \frac{\partial u}{\partial y} \right) dx. \tag{2}$$

Equation (2) may be repeated for the displacement v as

$$v_Q - v_P = \int_{x_P}^{x_Q} \left(\frac{\partial v}{\partial x} + m \frac{\partial v}{\partial y} \right) dx. \tag{3}$$

Replacing the partial derivatives by the strain components in that plane, we obtain

$$u_Q - u_P = \int_{x_P}^{x_Q} (\epsilon_x + m\alpha) dx \tag{4}$$

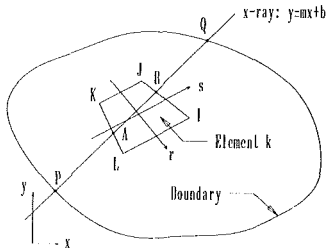


Fig. 1 Schematic representation of the problem under investigation

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16–19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 4, 1989; final revision, Aug. 2, 1990.

Paper No. 91-APM-14.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$v_Q - v_P = \int_{x_P}^{x_Q} (\beta + m\ \epsilon_y) dx \tag{5}$$

where $\gamma_{xy} = \alpha + \beta$. Since there exists no coupling between u and v in the formulation, we present the following development using u alone for the sake of brevity.

We now discretize the plane region into a number of four-noded elements. Then we have, from equation (4),

$$u_Q - u_P = \sum_k \int_{x_A}^{x_B} (\epsilon_x + m\alpha) dx, \tag{6}$$

where x_A and x_B are the intersections of the imaginary ray and the boundary edges of element k. Our goal at this stage is to obtain ϵ_x and α at element nodes. It suffices to express ϵ_x and α in terms of their nodal values. This may be fulfilled using interpolation technique commonly employed in isoparametric finite element formulation (see, for instance, the text by R. D. Cook et al., 1989). That is,

$$\begin{aligned} \epsilon_x &= N_i\ \epsilon_{xi}, \\ \alpha &= N_i\ \alpha_i,\ 1 \le i \le 4 \end{aligned} \tag{7}$$

where

$$N_1 = \frac{1}{4}(1+r)(1+s) \qquad N_2 = \frac{1}{4}(1-r)(1+s)$$
$$N_3 = \frac{1}{4}(1-r)(1-s) \qquad N_4 = \frac{1}{4}(1+r)(1-s).$$

Here, r and s are the natural coordinates in the transformed plane as shown in Fig. 2 for a generic element. In equation (7), ϵ_{xi} and α_i are ϵ_x and α evaluated at element node i, and indicial summation convention is used here and from this point on. The interpolation functions N_i are also used in coordinate transformation:

$$x = N_i x_i \text{ and } y = N_i y_i \tag{8}$$

where x_i and y_i are coordinate of node i. Substitution of (8) in (1) gives the X-ray equation in transformed plane:

$$C_0 + C_1 r + C_2 s + C_3 rs = 0 \tag{9}$$

where $C_0, \ldots C_3$ are constants. In general, C_3 does not vanish indicating that the imaginary ray in transformed plane is not straight.

Now, using equations (7) and (9) in equation (6), we obtain

$$u_Q - u_P = \sum_k \int_{r_A}^{r_B} N_i(\epsilon_{xi} + m\ \alpha_i) J\, dr \tag{10}$$

where

$$J = \left(\frac{\partial N_i}{\partial r} + \frac{\partial N_i}{\partial s}\frac{ds}{dr}\right) x_i.$$

Note that $dx = J\, dr$, and r_A and r_B are the natural coordinates of points A and B. This equation may be presented in matrix form as

$$u_Q - u_P = \sum_k \left\{ \int_{r_A}^{r_B} \mathbf{N}\, J\, dr\, \mathbf{E}^k \right\} \tag{11}$$

where the row matrix $\mathbf{N}$ and column matrix $\mathbf{E}^k$ for element k are

$$\mathbf{N} = [N_1 \quad mN_1 \ \ldots\ N_4 \quad mN_4]$$
$$(\mathbf{E}^k)^T = [\epsilon_{x1} \quad \alpha_1 \ \ldots\ \epsilon_{x4} \quad \alpha_4].$$

Note that s and ds/dr in equation (11) may be determined from equation (9) for a given r.

Since points P and Q are not necessarily the element nodes, interpolation between displacement u of neighboring nodes is employed to find u_P and u_Q in equation (11). We have found from our numerical experiments that accurate results can be achieved only when a midside node on a boundary edge is

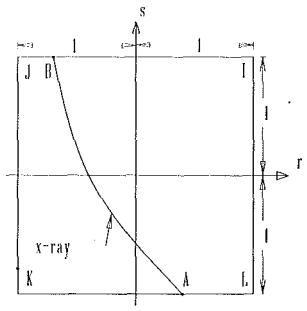


Fig. 2 Element and X-ray in transformed natural plane

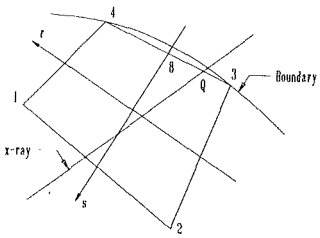


Fig. 3 A midside node of an element edge on boundary

inserted and quadratic interpolation is used. In making reference to Fig. 3, where corner nodes 3 and 4 and midside node 8 are on the boundary, we have

$$u_Q = \frac{1}{2}(r + r^2)u_4 + \frac{1}{2}(-r + r^2)u_3 + (1 - r^2)u_8 \tag{12}$$

where r is the natural coordinate and u_3, u_4, and u_8 are known displacements. The displacement u_P may be obtained following the same manner.

The integral in equation (11) generally cannot be integrated analytically. The Gauss quadrature is used here after a linear transformation is introduced so that the upper and lower limits become 1 and -1, respectively. The detail in this regard is given in the text by R.D. Cook et al. (1989).

Each imaginary ray will give us one equation of the form (11) which, after Gauss quadrature is evaluated, is written as follows

$$A_{ij}\ E_j = \Delta\ u_i,\ 1 \le j \le 2N \tag{13}$$

where A_{ij} represents the results from Gauss quadrature in equation (11), E_j designates all the unknowns in the model, $\Delta u_i = u_Q - u_P$ for X-ray i, and N is the total number of nodes in the model. Each node has two unknowns, ϵ_x and α; unknowns β and ϵ_y will be solved separately by replacing Δu_i in equation (13) by Δv_i. There are only a few elements in the model traversed by a ray, which results in only a few nonzero terms of A_{ij} in equation (13). It needs to be noted that, for a convex or a multiply-connected region, one ray could result in more than one equation of the form (13).

In order to solve the system, we need as many as $2N$ rays in many different injection angles or slopes. The system is posed in matrix form

$$\mathbf{AE} = \mathbf{\Delta}. \tag{14}$$

This system is likely to be ill-posed since the determinant of matrix $\mathbf{A}$ is either zero or vanishingly small. This difficulty may be overcome by increasing the number of rays to, say, twice as many. Then the least square technique is employed to eliminate the redundancy. That is, the system (14) is multiplied by the transpose of $\mathbf{A}$, now a rectangular matrix.

$$(\mathbf{A}^T\mathbf{A})\mathbf{E} = \mathbf{A}^T\mathbf{\Delta}$$

or

$$\mathbf{CE} = \mathbf{\Delta}' \tag{15}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

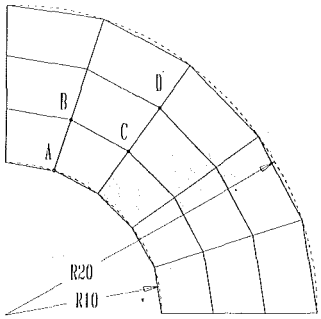


Fig. 4 A quarter model of the pressurized cylinder

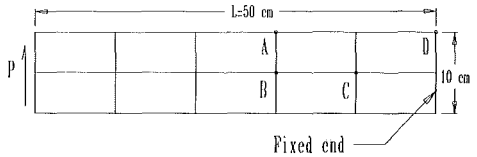


Fig. 5 Bending of a cantilever by an end force

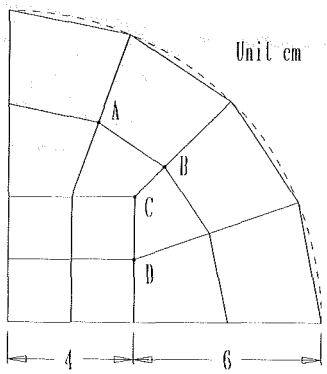


Fig. 6 A quarter model of a rotating disk

where $\mathbf{C}$ now is a $2N$ by $2N$ nonsymmetric matrix, and Δ' is a column matrix of compatible size. To also solve for the unknowns ϵ_y and β, we need only to augment the column matrix Δ by adding a column containing Δv_i. Consequently, Δ' is of size $2N \times 2$. Equation (15) can be solved by Gauss elimination scheme or other iterative scheme. It is often an important routine to check the least pivotal values during elimination scheme, which could serve as an indicator of sensitivity and accuracy.

Once the derivatives $\partial u/\partial x$ and $\partial u/\partial y$ are found, the displacements at internal nodes can be determined through integration over an edge with known displacement at one end. That is, referring back to Fig. 1,

$$u_J = u_I + \int_{x_J}^{x_I} \left(\frac{\partial u}{\partial x} + m \frac{\partial u}{\partial y} \right) dx \tag{16}$$

assuming that displacement u at node I is either given or found in the process. Here, the derivatives have to be expressed using linear interpolations between the derivatives solved in previous phase. Therefore, Gauss quadrature is employed again to evaluate

$$u_J = u_I + \int_{-1}^{1} \left\{ \frac{1}{2}(1+r) \left[\left(\frac{\partial u}{\partial x} \right)_I + m \left(\frac{\partial u}{\partial y} \right)_I \right] + \frac{1}{2}(1-r) \left[\left(\frac{\partial u}{\partial x} \right)_J + m \left(\frac{\partial u}{\partial y} \right)_J \right] \right\} \frac{x_J - x_I}{2} dr. \tag{17}$$

An interior node with unknown displacement could be common to several element edges along which the integration can be evaluated. In such an event, we average the results obtained to achieve better accuracy. Note that s is used in place of r in the above equation provided that integration is performed along an edge parallel to s-axis. Further calculation may be pursued to find stress information so long as material properties are provided.

Table 1 Comparison of calculated and analytical (in parenthesis) solutions at selected nodes in Fig. 4

	ϵ_x	ϵ_y	γ_{xy}	u	v
Node A	4.654 (4.377)	-3.373 (-2.977)	7.457 (7.354)	- -	- -
Node B	2.751 (2.768)	-1.445 (-1.368)	-4.237 (-4.137)	18.12 (18.50)	43.82 (44.65)
Node C	0.660 (0.7000)	0.6424 (0.7000)	-5.888 (-5.850)	32.98 (34.18)	33.84 (34.18)
Node D	0.7641 (0.7000)	0.7552 (0.7000)	-3.674 (-3.744)	30.06 (30.31)	31.39 (30.31)

Table 2 Comparison of calculated and analytical (in parenthesis) solutions at selected nodes in Fig. 5

	ϵ_x	ϵ_y	γ_{xy}	u	v
Node A	7.406 (7.200)	2.277 (2.160)	-76.04 (-76.44)	- -	- -
Node B	-0.019 (0.000)	0.026 (0.000)	-76.98 (-75.24)	0.055 (0.000)	416.6 (416.0)
Node C	-0.003 (0.000)	0.009 (0.000)	-41.91 (-41.64)	0.025 (0.000)	112.4 (112.0)
Node D	12.04 (12.00)	3.611 (3.600)	-0.094 (-0.360)	- -	- -

Table 3 Comparison of calculated and analytical (in parenthesis) solutions at selected nodes in Fig. 6

	ϵ_x	ϵ_y	γ_{xy}	u	v
Node A	22.07 (21.67)	14.41 (13.79)	-8.087 (-7.882)	63.69 (62.42)	150.8 (150.7)
Node B	17.97 (17.73)	18.25 (17.73)	-11.71 (-11.15)	116.6 (115.3)	115.6 (115.3)
Node C	21.53 (21.59)	21.56 (21.59)	-8.037 (-7.280)	102.9 (100.9)	101.2 (100.9)
Node D	23.03 (22.96)	26.25 (25.69)	-3.69 (-3.64)	109.1 (106.4)	53.81 (53.20)

3 Numerical Simulation

The proposed computational technique is applicable when boundary displacements are known *a priori*. It is understood that the boundary displacements are generally obtained by experimental means such as hologram interferometry. It is our goal here to report the feasibility of the proposed technique. No attempt has as yet been made to demonstrate the numerical simulation using experimental data. Only examples whose analytical solutions exist, are used in this paper.

The first example involves a thick-walled cylinder subject to internal pressure, p_i, whose radial displacement is given in many texts of elasticity (for example, Timoshenko and Goodier, 1970). Taking advantage of symmetry, only a quarter of the cross-section is used in the calculation. It is well understood that it will pose a good deal of difficulty to experimentally measure the displacement at inner wall if the inner radius gets too small. And it is unlikely to obtain experimental data along the imaginary cutting boundary. We use this example simply because it is one of the rare cases available which have analytical solutions in displacements. The boundary displacements calculated from the analytical solution simulate the experi-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

mental data. The 12-element mesh is shown in Fig. 4, which consists of 40 unknown derivatives. One hundred and five rays, ranging from +75 deg to −75 deg, are used in the calculation. The calculated strains and displacements together with their analytical counterparts at several selected nodes are reported in Table 1. Note that Poisson's ratio of 0.3 is used in calculating boundary displacements, and the values in Table 1 need to be multiplied by $p_i/3E$ to attain correct order of magnitude, where E is the modulus of elasticity. In general, accuracy in displacement is better than that in strains. This is because of the averaging mechanism used in finding displacements.

The second example considered is the bending of a cantilever, having a narrow rectangular cross-section of unit width, by an end force P. The analytical solutions to displacements u and v are given in the text by Timoshenko and Goodier (1970). The mesh (18 nodes, 10 elements) used in calculation is shown in Fig. 5, where the right end is fixed. We used 88 rays at 11 different angles ranging from +75 deg to −75 deg. Table 2 lists the tomographic result and the analytical solution. Note that $\nu = 0.3$ is used and values given in Table 2 need to be multiplied by 0.001 PL^3/EI, where $L = 50$ cm and EI is the flexural rigidity of the beam. Excellent agreement is attained, although fewer rays are used owing to the rather simple geometry.

A rotating disk of constant thickness is considered in the third example, whose exact displacement field is given in the text by Ugural and Fenster (1987). A quarter model, shown in Fig. 6, along with 80 rays at 17 different angles ranging from +80 deg to −80 deg, is used in calculating 38 unknown derivatives at 19 nodes. Comparison of the calculated result and exact solution is given in Table 3. Again, $\nu = 0.3$ is used and the multiplying factor for values in Table 3 is $\rho w^2/E$, where ρ is density, w angular speed, and E modulus of elasticity. Excellent agreement is again achieved.

4 Conclusion and Prospects

A computerized tomographic technique is developed to calculate strains and displacements at interior nodes from given boundary displacements regardless of material properties. In the framework of this new technique, the differential of displacements at two ends of a ray is evaluated by integrating the displacement derivatives along the X-ray. Research is now under way at Oakland to consider inversion of the two-dimensional Radon transform (see Vest, 1979) in place of direct integration method reported in this paper. Each ray emitted through the model results in one equation or more in case of convex region. To obtain satisfactory results, it is virtually necessary to use twice as many rays as the number of unknowns involved. According to our experience, for a given number of rays, the greater number of angles used to distribute rays, the more accurate result is attained. It needs intense study with regard to the possible effect on accuracy from element aspect ratio and higher-order elements. Excellent agreement between the results obtained numerically and analytically confirms the validity of the proposed tomographic technique.

Along this line of research our next phase will concentrate on incorporating the laboratory-measured displacements of a deformed three-dimensional object into the computational technique developed here. With the aid of a high-speed digital computer, the strain field of a selected transverse section may be pictorially displayed. Although known displacements around a boundary are used in the tomographic simulation, displacement measurements of a plane problem may be made along a designated region, particularly where direct optical wave exposure is prohibited. With these measurements as input, the strain field of the obscured region may then be shown. The application of tomographic technique to mechanics to our knowledge is as yet not available. As tomography has revolutionized medical diagnoses, we believe computational tomography will have a similar impact upon experimental mechanics.

References

Anderson, A. H., and Kak, A. C., 1982, "Digital Ray Tracing in Two-Dimensional Retractive Fields," *J. Acoust. Soc. Am.*, Vol. 72, pp. 1593–1598.

Cook, R. D., Malkus, D. S., and Plesha, M. E., 1989, *Concepts and Application of Finite Element Analysis*, John Wiley and Sons, New York.

Hennig, J., 1988, "Multiecho Image Sequences with Low Refocusing Flip Angles," *J. Mag. Res.*, Vol. 78, pp. 397–407.

Sweeney, D. W., and Vest, C. M., 1973, "Reconstruction of Three-Dimensional Refractive Index Fields from Multidirectional Interferometric Data," *Appl. Opt.*, Vol. 12, pp. 2649–2664.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York.

Ugural, A. C., and Fenster, S. K., 1987, *Advanced Strength and Applied Elasticity*, Elsevier, New York.

Vest, C. M., 1979, *Holographic Interferometry*, John Wiley and Sons, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A New Boundary Element Method for Three-Dimensional Coupled Problems of Consolidation and Thermoelasticity

G. F. Dargush
Assistant Professor,
Mem. ASME

P. K. Banerjee
Professor,
Mem. ASME

Department of Civil Engineering,
State University of New York at Buffalo,
Buffalo, NY 14260

A general boundary element method (BEM) for the complete three-dimensional Biot consolidation theory is developed which operates directly in the time domain and requires only boundary discretization. Consequently, the dimensionality of the problem is reduced by one and the method becomes quite attractive for geotechnical analyses, particularly those that involve extensive or infinite domains. Furthermore, as a result of a well-known analogy fully elaborated here, the new BEM formulation is equally applicable for quasi-static thermoelasticity. Several detailed examples are presented to illustrate the accuracy and suitability of this boundary element approach for both consolidation and thermomechanical analyses.

Introduction

Over the past decade, considerable attention has been directed toward developing a BEM for thermoelasticity and consolidation. The earliest work, done by Rizzo and Shippy (1977) and Cruse et al. (1977), pertained to a steady-state boundary-only formulation of thermoelasticity.

Banerjee and Butterfield (1981, 1982) presented a staggered procedure for solving the coupled equations. The algorithm requires the solutions of the transient pore fluid (or heat) flow equations followed by an elastic analysis including body forces at each time step. Unfortunately, this is not a boundary-only formulation, and complete volume discretization is required. Interestingly, this same scheme was applied by Kuroki et al. (1982) for two-dimensional consolidation, by Aramaki and Yasuhara (1985) for axisymmetric consolidation, and by Aramaki (1986) for consolidation problems which include thin layers with high permeability.

Another alternative involves casting the problem in the Laplace transform domain. Cheng and Liggett (1984) took that approach for two-dimensional quasi-static poroelasticity. While this is a boundary-only formulation, the numerical inversion of Laplace transform is neither efficient nor accurate for a complex problem. Moreover, it is impossible to extend such an algorithm to nonlinear problems. More recently, Nishimura (1985) developed the first two-dimensional time-domain boundary element formulation for regions infiltrated by an incompressible fluid. Numerical results for a single, one element problem were included. In a subsequent paper, Nishimura et al. (1986) presented two additional problems, but provided no further correlations with analytical results.

In a series of papers, Sladek and Sladek (1983, 1984a, 1984b) presented a collection of fundamental solutions, without any implementation in both Laplace transform and time domains, under the classifications of coupled, uncoupled, transient, and quasi-static thermoelasticity. Boundary integral equations were also included, although in several instances these were written inappropriately in terms of displacement and traction rates. Similar mathematical formulation papers have also been published by Cheng and Predeleanu (1987) and Masinda (1984). Meanwhile, Sharp and Crouch (1986) have directed efforts toward converting the thermal body force volume integral into a surface integral in a two-dimensional problem.

The present effort involves the development of boundary element methods for general three-dimensional Biot's theory of consolidation and the corresponding theory of coupled thermoelasticity. Included is capability to analyze arbitrarily shaped, multiregion bodies subjected to time-varying boundary conditions. The resulting BEM time-domain formulation and implementation represents the first of its kind for three-dimensional problems. The corresponding two-dimensional work was presented in Dargush (1987) and Dargush and Banerjee (1989a, b).

Governing Equations

Poroelasticity. Following Biot (1941), the total stress σ_{ij} is decomposed into effective stress σ'_{ij} of the soil and pore pressure p of the fluid that is,

$$\sigma_{ij} = \sigma'_{ij} - \delta_{ij}\beta p. \tag{1}$$

Then the effective stress is related to the strains via an isotropic, linear elastic constitutive law of the form

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 3, 1989; final revision, Sept. 15, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\sigma'_{ij} = \delta_{ij}\lambda\epsilon_{kk} + 2\mu\epsilon_{ij}. \tag{2}$$

The elastic constants, λ and μ, appearing in equation (2) are those obtained from soil samples under drained conditions. The dimensionless material parameter β in equation (1) can be defined as

$$\beta = 1 - \frac{K}{K'_s} = \frac{1}{B}\left[1 - \frac{K}{K_u}\right] \tag{3}$$

where $K = \lambda + 2\mu/3$ is the drained bulk modulus, B is Skempton's pore pressure coefficient, K_u the undrained bulk modulus and K'_s is an empirical constant which in certain circumstances equals the bulk modulus of the solid constituents (Rice and Cleary, 1976).

Equations (1) and (2) can be combined with the equilibrium and the strain displacement relations to produce Navier's equations generalized for a poroelastic body:

$$(\lambda + \mu)u_{j,ij} + \mu u_{i,jj} - \beta p_{,i} + f_i = 0. \tag{4}$$

The set of governing differential equations for poroelasticity must be supplemented by applying the principle of conservation of mass to the fluid:

$$q_{i,i} + \dot{m} = \bar{\psi} \tag{5}$$

where q_i is the fluid mass flux vector, m is fluid mass content per unit volume, and $\bar{\psi}$ is the fluid mass supply rate per unit volume.

Next, Darcy's law is invoked to relate flux to pore pressure gradient in a linear fashion. For an isotropic medium this becomes simply

$$q_i = \rho_o \kappa p_{,i} \tag{6}$$

with fluid mass density ρ_o and permeability κ. The fluid mass content must also be expressed in terms of displacements and pore pressure. A convenient form of this relationship, given by Rudnicki (1987), is

$$m - m_o = \beta\rho_o\left(u_{j,j} + \frac{\beta p}{\lambda_u - \lambda}\right) \tag{7}$$

in which m_o is the reference value of m and λ_u is the undrained Lamé modulus. Introducing equations (6) and (7) into (5) leads to the following expression for mass conservation:

$$\rho_o\kappa p_{,jj} - \frac{\rho_o\beta^2}{\lambda_u - \lambda}\dot{p} - \rho_o\beta\dot{u}_{j,j} + \bar{\psi} = 0. \tag{8}$$

After letting $\bar{\psi} = \rho_o\psi$ in equation (8), the complete coupled set of governing differential equations for poroelasticity can be rewritten as

$$(\lambda + \mu)u_{j,ij} + \mu u_{i,jj} - \beta p_{,i} + f_i = 0 \tag{9a}$$

$$\kappa p_{,jj} - \frac{\beta^2}{\lambda_u - \lambda}\dot{p} - \beta\dot{u}_{j,j} + \psi = 0. \tag{9b}$$

The material parameters present in equation (9) are not necessarily those typically measured for soils, and other choices are perfectly valid (Dargush, 1987; Dargush and Banerjee, 1989a, d).

An interesting reformulation of equation (9) in terms of the fluid mass content leads to the following set of governing differential equations:

$$(\lambda_u + \mu)u_{j,ij} + \mu u_{i,jj} - \left(\frac{\lambda_u - \lambda}{\beta}\right)m_{,i} = 0 \tag{10a}$$

$$\left[\kappa\left(\frac{\lambda_u - \lambda}{\beta^2}\right)\frac{\lambda + 2\mu}{\lambda_u + 2\mu}\right]m_{,jj} = \dot{m}. \tag{10b}$$

Now, upon examination of equation (10b) it is evident that in poroelasticity, as emphasized by Rice and Cleary (1976), the fluid mass content must satisfy a simple diffusion equation. The diffusivity, c, is a material characteristic, which becomes

$$c = \kappa\left(\frac{\lambda_u - \lambda}{\beta^2}\right)\frac{\lambda + 2\mu}{\lambda_u + 2\mu}. \tag{11}$$

Thermoelasticity. A complete derivation of thermoelastic theory can be found in the textbooks by Boley and Weiner (1960) and Nowacki (1986). In particular, the classical theory assumes infinitesimal deformations and linear isotropic materials. Thus, the linearized strain-displacement relationship is once again employed, along with the constitutive laws of Duhamel-Neumann

$$\sigma_{ij} = \delta_{ij}\lambda\epsilon_{kk} + 2\mu\epsilon_{ij} - \delta_{ij}(3\lambda + 2\mu)\alpha(\theta - \theta_o) \tag{12}$$

and Fourier

$$q_i = -k\theta_{,i}. \tag{13}$$

In equations (12) and (13) θ is the present temperature, while θ_o represents the temperature of a zero-stress reference state, and q_i is the heat flux vector. Additionally, λ and μ are the usual Lamé isothermal elastic constants, α is the coefficient of thermal expansion, and k represents the isotropic thermal conductivity.

Then, application of the laws of conservation of momentum and energy lead, in the absence of inertia, to the following set of equations

$$(\lambda + \mu)u_{j,ij} + \mu u_{i,jj} - (3\lambda + 2\mu)\alpha\theta_{,i} + f_i = 0 \tag{14a}$$

$$k\theta_{,jj} - \rho c_\epsilon\dot{\theta} - (3\lambda + 2\mu)\alpha\theta_o\dot{u}_{j,j} + \psi = 0, \tag{14b}$$

in which ρ is the mass density and c_ϵ is the specific heat at constant deformation, while f_i and ψ are the body forces and sources, respectively. The theory portrayed by equation (14) is formally named coupled quasi-static thermoelasticity (CQT). In the present work the general, fully coupled theory governed by equation (14) will be retained because it is the only theory that provides the exact thermodynamic analogy between the poroelasticity and thermoelasticity. The more efficient and practical uncoupled approach is detailed in Dargush (1987), Dargush and Banerjee (1989b, 1989c).

The governing differential equations (14) can be reformulated by employing the entropy density, s, rather than the temperature as an explicit variable. The result is

$$\left(\lambda + \frac{\theta_o\beta^2}{\rho c_\epsilon} + \mu\right)u_{j,ij} + \mu u_{i,jj} - \left(\frac{\theta_o\beta}{\rho c_\epsilon}\right)s_{,i} = 0 \tag{15a}$$

$$\left[\frac{k}{\rho c_\epsilon}\frac{(\lambda + 2\mu)}{\left(\lambda + \frac{\theta_o\beta^2}{\rho c_\epsilon} + 2\mu\right)}\right]s_{,jj} = \dot{s} \tag{15b}$$

where, for simplicity, $\beta = (3\lambda + 2\mu)\alpha$.

Interestingly, equation (15b) suggests that molecular disorder, quantified by entropy density, diffuses through a CQT body. The rate at which this process progresses at any time depends upon the existing entropy density distribution and a material parameter c. This thermal diffusivity is defined by

$$c = \left(\frac{k}{\rho c_\epsilon}\right)(\lambda + 2\mu)/\left(\lambda + \frac{\theta_o\beta^2}{\rho c_\epsilon} + 2\mu\right). \tag{16}$$

Poroelastic-Thermoelastic Analogy. From the presentations of the previous two sections, it has become evident that there is a strong analogy between poroelasticity and thermoelasticity. It appears that this analogy was first discovered by Biot (1956), who earlier pioneered much of the work in poroelasticity. One can see the similar roles played by the pore pressure and the temperature. However, the correspondence runs much deeper (as developed in the present work) since the two phenomena are completely analogous.

Many useful relationships can be established by conducting a term-by-term investigation of equations (10) versus (15). For

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Summary of the analogy

Thermoelasticity	correspondence		Poroelasticity
entropy density	s	m	fluid mass content per unit volume
temperature	θ	p	pore pressure
heat flux	q	q	fluid mass flux
diffusivity	c	c	diffusivity
$c=\dfrac{k}{\rho c_\epsilon}\dfrac{\lambda+2\mu}{\lambda_s+2\mu}$			$c=\kappa\left(\dfrac{\lambda_u-\lambda}{\beta^2}\right)\dfrac{\lambda+2\mu}{\lambda_u+2\mu}$
isentropic Lamé constant	λ_s	λ_u	undrained Lamé constant
$\lambda_s=\lambda+\dfrac{\theta_o\beta^2}{\rho c_\epsilon}$			$\lambda_u=\lambda+\dfrac{B\beta K}{1-B\beta}$
thermoelastic constant	β	β	poroelastic constant
$\beta=(3\lambda+2\mu)\alpha$			$\beta=\dfrac{3}{B}\left[\dfrac{\lambda_u-\lambda}{3\lambda_u+2\mu}\right]$
conductivity over reference temperature	$\dfrac{k}{\theta_u}$	κ	permeability

example, consider the coefficients of $u_{j,ij}$. In this case, the terms $\lambda+\theta_o\beta^2/\rho c_\epsilon$ in CQT correspond to the undrained Lamé modulus λ_u of poroelasticity. Actually this is not unexpected, since under the enforcement of isentropic deformation the relevant Lamé constant, say λ_s, becomes equivalent to $\lambda+\theta_o\beta^2/\rho c_\epsilon$. Thus isentropic deformation in CQT is the counterpart of undrained deformation in poroelasticity.

Table 1 provides a detailed summary of the complete analogy. In the present work, this analogy is exploited to the fullest, so that the boundary element formulation is directly applicable to both poro and thermoelasticity. However, as a matter of convenience, poroelastic terminology will be used throughout the remainder of this paper, except for the thermoelastic example.

Boundary Integral Formulation

A reciprocal theorem generally provides a convenient starting point for a direct boundary element formulation. For the coupled problem at hand, it appears that Ionescu-Cazimir (1964) was the first to state explicitly an appropriate reciprocal theorem, although certainly the groundwork was laid earlier by Biot (1959). In the context of quasi-static poroelasticity that theorem can be written for a body of volume V and surface S with zero initial conditions, in the following time-domain form:

$$\int_S[\dot{t}_i^{(1)}*u_i^{(2)}+q^{(1)}*p^{(2)}-\dot{u}_i^{(1)}*t_i^{(2)}-p^{(1)}*q^{(2)}]dS + \int_V[\dot{f}_i^{(1)}*u_i^{(2)}+\psi^{(1)}*p^{(2)}-\dot{u}_i^{(1)}*f_i^{(2)}-p^{(1)}*\psi^{(2)}]dV=0 \tag{17}$$

where $q=q_in_i$ is the normal flux. The superscripts $^{(1)}$ and $^{(2)}$ denote any two independent states of displacement rates, traction rates, pore pressure, flux, body force rate, and body source which satisfy the governing differential equations, in the body defined by $[\dot{u}_i^{(1)}, \dot{t}_i^{(1)}, \dot{p}^{(1)}, q^{(1)}, \dot{f}_i^{(1)}, \psi^{(1)}]$ and $[u_i^{(2)}, t_i^{(2)}, p^{(2)}, q^{(2)}, f_i^{(2)}, \psi^{(2)}]$, respectively. The symbol $*$ indicates a Riemann convolution integral, where, for example,

$$q^{(1)}*p^{(2)}=\int_0^t q^{(1)}(X, t-\tau)p^{(2)}(X, \tau)d\tau = \int_0^t q^{(1)}(X, \tau)p^{(2)}(X, t-\tau)d\tau. \tag{18}$$

It should be mentioned that the form of equation (17) above will lead to a boundary element formulation with displacement, traction, pore pressure, and flux as primary quantities. Other researchers, notably Predeleanu (1981) and Sládek and Sládek (1984b), have written equally valid reciprocal statements that, unfortunately, lead to a much less desirable set of primary variables requiring boundary conditions in terms of displacement and traction rates.

This reciprocal theorem involves any two independent states existing in the body of interest. Let one of those states, say state (2), be the, as yet, unknown solution to a given boundary value problem. Then the remaining state may be chosen arbitrarily. However, by initially selecting for state (1), the infinite space fundamental solution to a unit step force in the j-direction acting at ξ within V (Nowacki, 1966, 1986; Cleary, 1977; Rudnicki, 1987) and beginning at time zero, the volume integrals in the reciprocal theorem can be made to vanish (Banerjee and Butterfield, 1981).

Substitution of the point force solution in equation (17) and specifying zero body forces and sources leads to the integral representation:

$$\delta_{ij}u_i(\xi, t)=\int_S[\dot{G}_{ij}(X-\xi, t)*t_i(X, t)+\dot{G}_{pj}(X-\xi, t)*q(X, t) - \dot{F}_{ij}(X-\xi, t)*u_i(X, t)-\dot{F}_{pj}(X-\xi, t)*p(X, t)]dS(X). \tag{19}$$

This is an integral equation for interior displacement that involves only boundary quantities. Therefore, volume integration has indeed been eliminated through the use of an infinite space Green's function. In equation (19), the functions G_{ij} and G_{pj} (not G_{pj}) are the fundamental solutions for displacement and pore pressure, whereas F_{ij} and F_{pj} (not F_{pj}) are the corresponding surface traction and normal flux.

A similar relationship is also desired for interior pore pressure. To that end, return to the reciprocal theorem (17) and select, instead, for state (1) the infinite space response to a unit pulse fluid mass source, acting at time zero and at point ξ within V (Nowacki, 1966, 1986; Cleary, 1977, Rudnicki, 1987). Substitution of the point source solution into the reciprocal theorem (17) and assuming no specified body forces and sources leads to:

$$p(\xi, t)=\int_S[\dot{G}_{ip}(X-\xi, t)*t_i(X, t)+\dot{G}_{pp}(X-\xi, t)*q(X, t) - \dot{F}_{ip}(X-\xi, t)*u_i(X, t)-\dot{F}_{pp}(X-\xi, t)*p(X, t)]dS(X). \tag{20}$$

This, of course, is the desired statement for interior pore pressure in terms of boundary quantities. The functions G_{ip}, G_{pp}, F_{ip}, and F_{pp} are the displacements, pore pressure, surface tractions, and normal flux, respectively, derived from the fundamental unit source solutions.

Equations (19) and (20) can now be combined and rewritten in a more convenient matrix notation as

$$\begin{Bmatrix} u_j(\xi, t) \\ p(\xi, t) \end{Bmatrix} = \int_S \left(\begin{bmatrix} \dot{G}_{ij} & \dot{G}_{ip} \\ \dot{G}_{pj} & \dot{G}_{pp} \end{bmatrix}^T * \begin{Bmatrix} t_i(X, t) \\ q(X, t) \end{Bmatrix} - \begin{bmatrix} \dot{F}_{ij} & \dot{F}_{ip} \\ \dot{F}_{pj} & \dot{F}_{pp} \end{bmatrix}^T * \begin{Bmatrix} u_i(X, t) \\ p(X, t) \end{Bmatrix} \right) dS(X). \tag{21}$$

However, this can even be compacted further by generalizing the displacement and traction vectors to include pore pressure and flux, respectively, as an additional component. Thus, let

$$u_\alpha=\{u_1\ u_2\ u_3\ p\}^T \tag{22a}$$

$$t_\alpha=\{t_1\ t_2\ t_3\ q\}^T, \tag{22b}$$

where the Greek index α, and subsequently β, varies from one to four. Then (21) becomes simply

$$u_\alpha(\xi, t)=\int_S[\dot{G}_{\beta\alpha}*t_\beta(X, t)-\dot{F}_{\beta\alpha}*u_\beta(X, t)]dS(X). \tag{23}$$

Equation (23) can be viewed as a generalized Somigliana's identity for quasi-static poroelasticity or thermoelasticity and, as such, is an exact statement for the interior displacements and pore pressure at any point ξ within V at any time t. However, to determine those interior quantities, the entire

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

history of the boundary values of u_α and t_α must be known. In a well-posed boundary value problem only half of that information is given at each instant of time. In order to obtain the missing information, and in essence solve the problem, the point ξ must be moved to the boundary. Then equation (23) can be expressed as

$$c_{\beta\alpha}(\xi)u_\beta(\xi, t) = \int_s [\dot{G}_{\beta\alpha} * t_\beta(X, t) - \dot{F}_{\beta\alpha} * u_\beta(X, t)]dS(X) \quad (24)$$

where $c_{\beta\alpha}(\xi)$ is a matrix of constants dependent only upon the local geometry of the boundary at ξ. For ξ on a smooth surface, $c_{\beta\alpha}$ reduces to $\delta_{\beta\alpha}/2$.

Numerical Implementation

The boundary integral equation (24) is an exact statement. No approximations have been introduced other than those used to formulate the boundary value problem. However, in order to apply equation (24) for the solution of practical engineering problems, approximations are required in both time and space.

For the temporal discretization, the time interval from zero to t is divided into N equal increments of duration Δt. Within each time increment, the primary field variables, t_β and u_β, are assumed constant. As a result, these quantities can be brought outside of the time integral. Since the integrand remaining is known in explicit form from the fundamental solutions, the required temporal integration can be performed analytically, and written as

$$G_{\beta\alpha}^{N+1-n}(X-\xi) = \int_{(n-1)\Delta t}^{n\Delta t} \dot{G}_{\beta\alpha}(X-\xi, t-\tau)d\tau. \quad (25)$$

Combining this, and similar expressions for the $F_{\beta\alpha}$ integral, with equation (24) produces

$$c_{\beta\alpha}(\xi)u_\beta^N(\xi) = \sum_{n=1}^{N} \int_s \Big[G_{\beta\alpha}^{N+1-n}(X-\xi)t_\beta^n(X) - F_{\beta\alpha}^{N+1-n}(X-\xi)u_\beta^n(X) \Big] dS(X). \quad (26)$$

Details of the integrated kernel functions appearing in equation (26) can be found in the Appendix.

Next, spatial discretization is introduced in order to evaluate the surface integrals. In the present implementation, both linear and quadratic boundary elements are available for the description of the geometry as well as the primary field variables. Once this is accomplished, the nodal generalized displacements and tractions are brought outside the surface integral and the remaining shape function-kernel products are integrated numerically. Sophisticated, self-adaptive integration algorithms are employed to ensure accuracy and numerical efficiency (Banerjee et al., 1988).

With the discretization of the boundary integral equation in both time and space complete, a system of algebraic equations can be developed to permit the approximate solution of the original quasi-static problem. This is accomplished by systematically writing the integral equations at each global boundary node. The nodal collocation process produces a global set of equations of the form

$$\sum_{n=1}^{N} ([G^{N+1-n}]\{t^n\} - [F^{N+1-n}]\{u^n\}) = \{0\}, \quad (27)$$

in which $\{t^n\}$ and $\{u^n\}$ are nodal quantities with the superscript referencing the time-step index. It should be noted that during this collocation process, the indirect "rigid body" technique for transient problems (Banerjee and Ahmad, 1985) is employed to determine the strongly singular diagonal block of $[F^1]$.

In a well-posed problem, at any time t, the set of global generalized nodal displacements and tractions will contain exactly $4P$ unknown components, where P is the total number of functional nodes. Then, as the final stage in the assembly process, equation (27) can be rearranged to form (Banerjee and Ahmad, 1985; Dargush, 1987; Dargush and Banerjee, 1989a, 1989b):

$$[A^1]\{x^N\} = [B^1]\{y^N\} - \sum_{n=1}^{N} ([G^{N+1-n}]\{t^n\} - [F^{N+1-n}]\{u^n\}) \quad (28)$$

in which $\{x^N\}$ and $\{y^N\}$ represent the unknown and known nodal components, respectively. In addition, the summation represents the effect of past events. Thus, all quantities on the right-hand side of equation (28) are known at time step N. For the general purpose numerical implementation (GP-BEST), the formulation of equation (28) is extended to incorporate multiregion capability with time-dependent boundary conditions including springs, surface convection, and thermally resistant interfaces.

It may appear at first glance that in equation (28) a prodigious amount of numerical calculations are involved in the

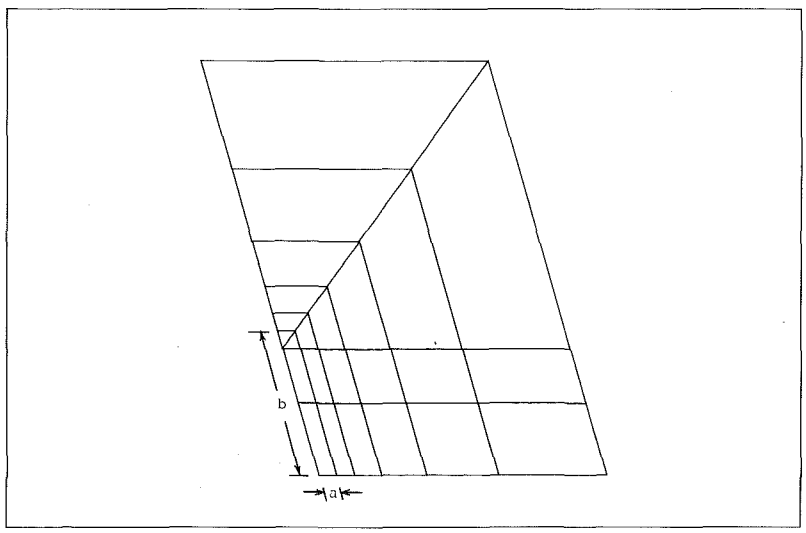


Fig. 1

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

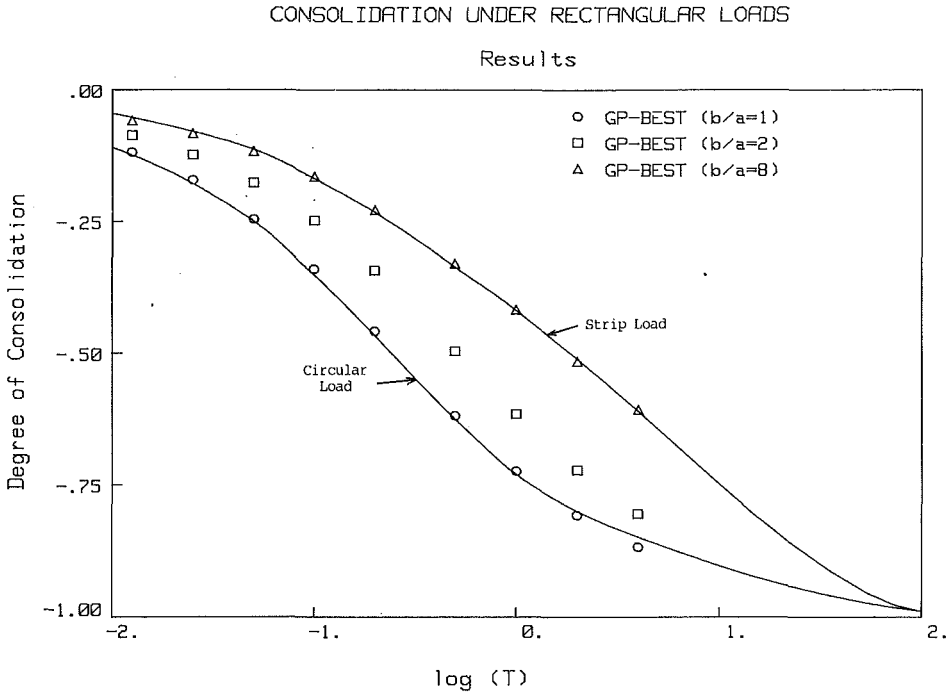


Fig. 2

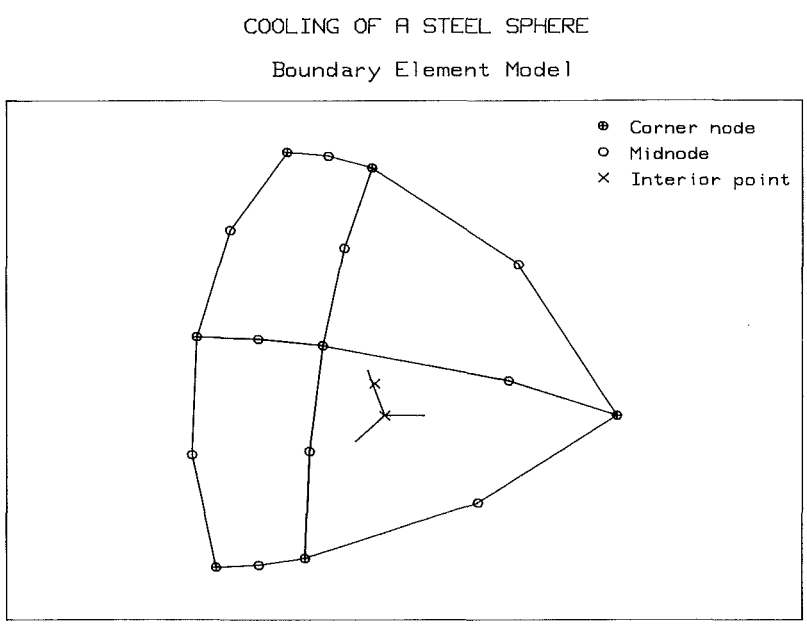


Fig. 3

solution process. However, a careful examination by substituting $n = 1, 2, \ldots$ in equation (28) will reveal that for each time step a new set of G and F matrices, representing the influence of the first time step on the current time values needs to be calculated. Since these involve very well-behaved integrals, the influence of which diminishes as the time stepping advances, the cost of computation for them is only a fraction of that involved in the leading time matrices A^1 and B^1, which are calculated only once. Typically if the cost for the first time step is unity, that for each of the second and third time step would be about 0.5, and for subsequent time steps the cost of each solution is reduced to about 0.25. This reduction is only possible because the precision with which the adaptive integration (Banerjee et al., 1988) needs to be carried out can be diminished with time without much loss of accuracy. As a result at the subelement level the order of integration can be progressively reduced from an initial order of 4×4 to 3×3 and finally to 2×2. It is important also to note that while the CPU requirements can be reduced this way, little can be done about economizing on disk storage requirements.

Applications

Consolidation Under Rectangular Loads. This poroelastic application involves the fully three-dimensional problem of the consolidation of a fully saturated half-space under abruptly applied and maintained rectangular loads of various dimensions. Outside the loaded area a traction-free boundary with zero pore pressure is assumed. The three-specific cases considered involve length (b) to width (a) ratios of 1.0, 2.0, and 8.0 for the loaded area. A boundary element model corresponding to the latter case is shown in Fig. 1. Notice that with the present formulation, discretization is required on only a

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

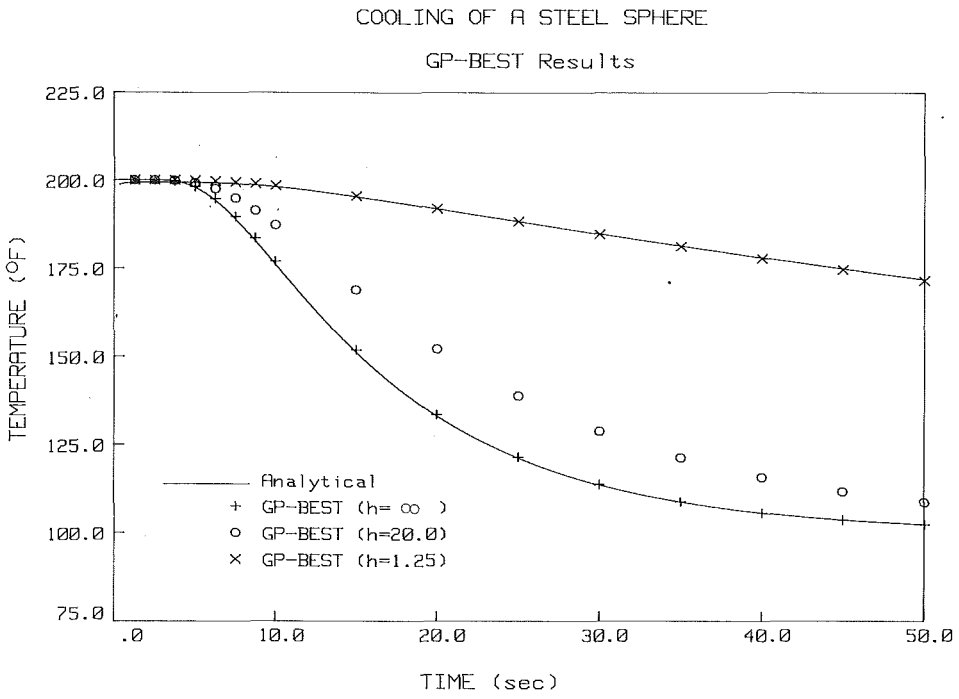


Fig. 4

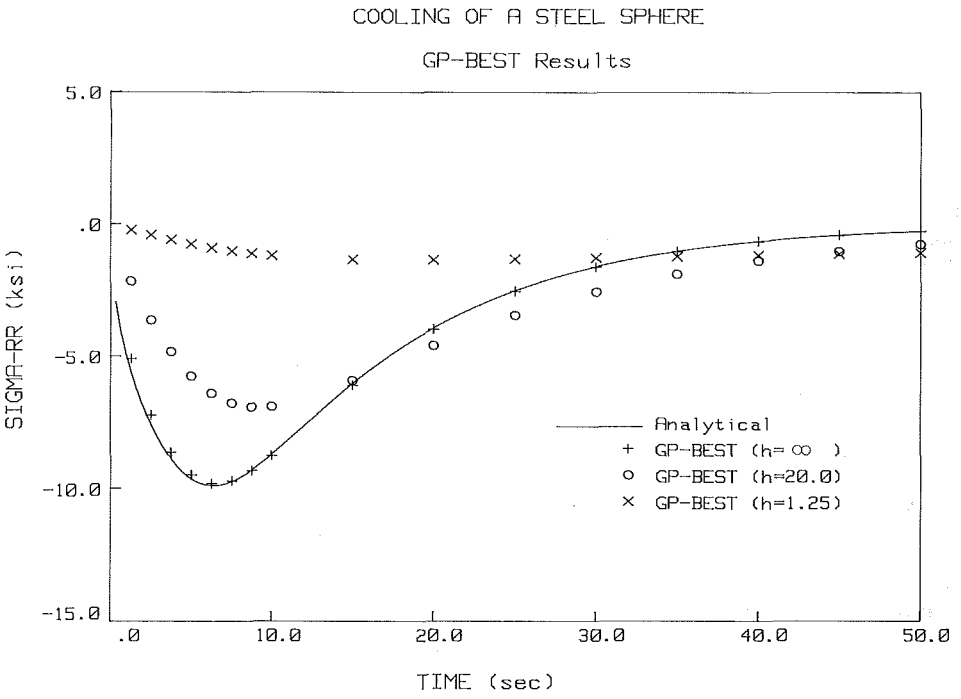


Fig. 5

small portion of the half-space because the extensive mesh over farther regions has negligible influence in the area of interest (Banerjee and Butterfield, 1981). Additionally, quadrantal symmetry is utilized to reduce computational effort. The material properties were arbitrarily selected as $E=1.0$, $\nu=0.0$, $K/\gamma_\omega=1.0$, $B=1.0$ and $\nu_u=0.5$, corresponding to an incompressible fluid constituent.

Figure 2 displays the results for the degree of consolidation at the center of the loaded area versus nondimensional time. For comparative purposes, the analytical curves of McNamee and Gibson (1960) for consolidation under strip and circular loads are also shown. It is evident, from the graph, that the behavior under a rectangular load with an aspect ratio of eight is nearly identical to the response due to a strip load. At the other end of the spectrum, the results for a square load closely follow the analytical solution for a circular load, with slight deviation present at large times. As expected, the boundary element results for $b/a=2$ fall between the two limiting cases. The CPU time for a 20 time-step analysis for each case was about 60 minutes on a SUN 4/110 desk-top computer.

Cooling of a Steel Sphere. As an application in thermoelasticity, consider a solid sphere of radius 1.5 in., initially at a uniform temperature of 200°F, cooled in 100°F air by convection. The thermomechanical properties of the steel are assumed as follows:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

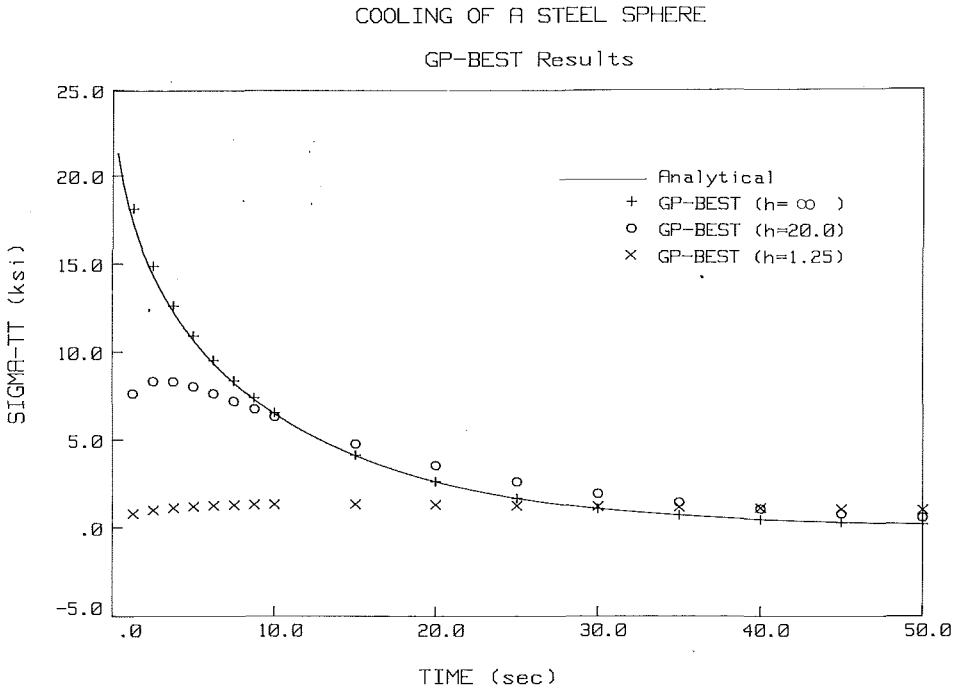


Fig. 6

$E = 30.0 \times 10^6$ psi $\quad \rho c_\epsilon = 283$ in.-lb./in.^{3} °F
$\nu = 0.30$ $\quad k = 5.8$ in.-lb./sec. in. °F
$\alpha = 6.0 \times 10^{-6}$in./in. °F $\quad c = 0.0203$ in.^{2}/sec.

Three levels of convection (characterized by the film coefficient h (Boley and Weiner, 1960)) are examined: very slow cooling ($h = 1.25$ in.-lb./sec. in.^{2} °F), moderate cooling ($h = 20.0$ in.-lb./sec. in.^{2} °F), and rapid cooling ($h \to \infty$). Analytical solutions for the temperature distribution in the sphere can be found in Carslaw and Jaeger (1947), while the corresponding stress fields for the rapid cooling case are presented in Nowacki (1986).

The GP-BEST boundary element model for this problem consists of two eight-noded and two six-noded surface elements on an octant of the sphere. As a result, a total of 17 source points are utilized, along with the octahedral symmetry option as shown in Fig. 3. A time step of 1.25 sec. is selected for all of the analyses.

The resulting temperatures versus time for the point at the center of the sphere are shown in Fig. 4. It is apparent that an excellent correlation is obtained for slow as well as rapid cooling. Meanwhile, Fig. 5 displays the radial stresses at that point. The correlation is again very good throughout the time history. Notice that, as expected, with slower surface cooling, the peak stress magnitude decreases while the time to reach steady-state (i.e., zero stress) increases. Finally, in Fig. 6, the tangential surface stresses are plotted. Even in the severe, rapid cooling case, the GP-BEST results are quite accurate. In fact, this highlights one of the primary advantages of the method for transient thermoelastic problems. The steep thermal gradients that are present at the initial instant can be captured with a high degree of precision. Consequently, the thermal surface stresses can also be calculated accurately. It is of interest to note that these results are within 0.1 percent of those of the corresponding uncoupled thermoelastic problem (Dargush and Banerjee, 1989c), indicating the well-known fact that for all linear thermoelastic problems without inertia effects, the uncoupled theory is perfectly satisfactory. The CPU time for a 40 time-step analysis for this fully coupled thermoelastic analysis was about 25 min on a SUN 4/110 desk-top computer.

Conclusions

A boundary element method has been developed for application to three-dimensional problems of poroelasticity and coupled thermoelasticity. The method is formulated directly in the time domain and requires only surface discretization. Thus, the dimensionality of the problem is reduced by one. This is particularly attractive for poroelasticity, since most practical problems involve large or semi-infinite domains. In addition, for thermoelasticity, the BEM is able to accurately capture steep temperature gradients and surface stresses. Two simple numerical examples are included to validate this new formulation. The major disadvantage of the analysis appears to be its relatively large data storage requirements as time stepping progresses. This, however, in the opinion of the authors, is more than compensated by its superior accuracy.

References

Aramaki, G., 1986, "Boundary Elements for Thin Layers With High Permeability in Biot's Consolidation Analysis," *Appl. Math. Modelling*, Vol. 10, pp. 82–86.

Aramaki, G., and Yasuhara, K., 1985, "Application of the Boundary Element Method for Axisymmetric Biot's Consolidation," *Engrg. Anal.*, Vol. 2, pp. 184–191.

Banerjee, P. K., and Ahmad, S., 1985, "Advanced Three-Dimensional Dynamic Analysis by Boundary Element Methods," *Advanced Topics in Boundary Element Analysis*, T. A. Cruse, A. B. Pifko, and H. Armen, eds., AMD-V72, ASME, New York.

Banerjee, P. K., and Butterfield, R., 1981, *Boundary Element Methods in Engineering Science*, McGraw-Hill, London.

Banerjee, P. K, and Butterfield, R., 1982, "Transient Flow Through Porous Elastic Media," *Developments in BEM*, Vol. 2, P. K. Banerjee and R. P. Shaw, eds., Elsevier Applied Science Publishers, London, pp. 19–35.

Banerjee, P. K., Wilson, R. B., and Miller, N., 1988, "Advanced Elastic and Inelastic Three-Dimensional Analysis of Gas Turbine Engine Structures by BEM," *Int. Jour. Num. Methods in Engrg.*, Vol. 26, pp. 393–411.

Biot, M. A., 1941, "General Theory of Three-Dimensional Consolidation," *J. Appl. Phys.*, Vol. 12, pp. 155–164.

Biot, M. A., 1956, "Thermoelasticity and Irreversible Thermodynamics," *J. Appl. Phys..*, Vol. 27, pp. 240–253.

Biot, M. A., 1959, "New Thermomechanical Reciprocity Relations with Application to Thermal Stress Analysis," *Journal of the Aerospace Sciences*, Vol. 26, No. 7, pp. 401–408.

Boley, B. A., and Weiner, J. H., 1960, *Theory of Thermal Stresses*, John Wiley and Sons, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Carslaw, H. S., and Jaeger, J. C., 1947, *Conduction of Heat in Solids*, Claredon Press, Oxford, U.K.

Cheng, A. H-D., and Liggett, J. A., 1984, "Boundary Integral Equation Method for Linear Porous-Elasticity with Applications to Soil Consolidation," *Int. J. Num. Meth. Engrg.*, Vol. 20, pp. 255–278.

Cheng, A., and Predeleanu, M., 1987, "Transient Boundary Element Formulation for Linear Poroelasticity," *Appl. Math. Modeling*, Vol. 11, pp. 285–290.

Cleary, M. P., 1977, "Fundamental Solutions for a Fluid-Saturated Porous Solid," *Int. J. Solids Structures*, Vol. 13, pp. 785–806.

Cruse, T. A., Snow, D. W., and Wilson, R. B., 1977, "Numerical Solutions in Axisymmetric Elasticity," *Comp. and Struct.*, Vol. 7, pp. 445–451.

Dargush, G. F., 1987, "Boundary Element Methods for the Analgous Problems of Thermomechanics and Consolidation," Ph.D. Dissertation, State University of New York at Buffalo, Buffalo, New York.

Dargush, G. F., and Banerjee, P. K., 1989a, "A Time Domain Boundary Element Method for Poroelasticity," *Int. J. Num. Meth. Engrg.*, Vol. 28, pp. 2423–2449.

Dargush, G. F., and Banerjee, P. K., 1989b, "Development of a Boundary Element Method for Time-Dependent Planar Thermoelasticity," *Int. J. Solids Struct.*, Vol. 25, pp. 999–1021.

Dargush, G. F., and Banerjee, P. K., 1989c, "Boundary Element Methods in Three-Dimensional Thermoelasticity," *Int. J. Solids Struct.*, Vol. 26, pp. 199–216.

Dargush, G. F., and Banerjee, P. K., 1989d, "Boundary Element Method for Poroelastic and Thermoelastic Analysis," *Developments in BEM-5*, P. K. Banerjee and R. B. Wilson, eds., Elsevier Applied Science, London, pp. 113–157.

Ionescu-Cazimir, V., 1964, "Problem of Linear Coupled Thermoelasticity. Theorems on Reciprocity for the Dynamic Problem of Coupled Thermoelasticity. I," *Bulletin de l'académie Polonaise des Sciences, Series des sciences techniques*, Vol. 12, No. 9, pp. 473–488.

Kuroki, T., Ito, T., and Onishi, K., 1982, "Boundary Element Method in Biot's Linear Consolidation," *Appl. Math. Modelling*, Vol. 6, pp. 105–110.

Masinda, J., 1984, "Application of the Boundary Element Method to 3D Problems of Non-stationary Thermoelasticity," *Engrg. Anal.*, Vol. 1, pp. 66–69.

McNamee, J., and Gibson, R. E., 1960, "Plane Strain and Axially Symmetric Problems of Consolidation of a Semi-infinite Clay Stratum," *Quart. Jour. Mech. and Appl. Math.*, Vol. 13, pp. 210–227.

Nishimura, N., 1985, "A BIE Formulation for Consolidation Problems," *Boundary Elements VII, Proceedings of the 7th International Conf.*, C. A. Brebbia and G. Maier, eds., Springer-Verlag, Berlin.

Nishimura, N., Umeda, A., and Kobayashi, S., 1986, "Analysis of Consolidation by Boundary Integral Equation Method," *Boundary Elements, Proceedings of the International Conference*, Beijing, DuQuinghua, ed., pp. 119–126.

Nowacki, W., 1966, "Green's Functions for a Thermoelastic Medium (Quasistatic Problem)," *Bull. Inst. Polit. Jasi, Sierie noua*, Vol. 12, Nos. 3–4, pp. 83–92.

Nowacki, W., 1986, *Thermoelasticity*, Pergamon, Warsaw.

Predeleanu, M., 1981, "Boundary Integral Method for Porous Media, *Boundary Element Methods*, C. A. Brebbia, ed., Springer, New York, pp. 325–334.

Rice, J. R., and Cleary, M. P., 1976, "Some Basic Stress Diffusion Solutions for Fluid-Saturated Elastic Porous Media with Compressible Constituents," *Reviews of Geophysics and Space Physics*, Vol. 14, No. 2, pp. 227–241.

Rizzo, F. J., and Shippy, D. J., 1977, "An Advanced Boundary Integral Equation Method for Three-Dimensional Thermoelasticity," *Int. J. Num. Meth. Engrg.*, Vol. 11, pp. 1753–1768.

Rudnicki, J. W., 1987, "Fluid Mass Sources and Point Forces in Linear Elastic Diffusive Solids," *Mechanics of Materials*, Vol. 5, pp. 383–393.

Sharp, S., and Crouch, S. L., 1986, "Boundary Integral Methods for Thermoelasticity Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 298–302.

Sladek, V., and Sladek, J., 1983, "Boundary Integral Equation Method in Thermoelasticity Part I: General Analysis," *Appl. Math. Modelling*, Vol. 7, pp. 241–253.

Sladek, V., and Sladek, J., 1984a, "Boundary Integral Equation Method in Thermoelasticity Part III: Uncoupled Thermoelasticity," *Appl. Math. Modelling*, Vol. 8, pp. 413–418.

Sladek, V., and Sladek, J., 1984b, "Boundary Integral Equation Method in Two-Dimensional Thermoelasticity," *Engrg. Anal.*, Vol. 1, pp. 135–148.

APPENDIX

Three-dimensional kernels are provided, based upon continuous source and fundamental solutions. As a result, the following relationships must be used to determine the proper form of the functions required in equation (26) for the boundary element discretization. That is,

$$G^n_{\alpha\beta}(X-\xi) = G_{\alpha\beta}(X-\xi,\ n\Delta t) \qquad \text{for } n=1$$
$$G^n_{\alpha\beta}(X-\xi) = G_{\alpha\beta}(X-\xi, n\Delta t) - G_{\alpha\beta}(X-\xi,\ (n-1)\Delta t) \qquad \text{for } n>1.$$

In the specification of these kernels below, the arguments $(X-\xi,\ t)$ are assumed and thermoelastic material parameters are used, the corresponding poroelastic parameters can be substituted from Table 1.

x_i coordinates of integration point

ξ_i coordinates of field point

$y_i = x_i - \xi_i,\ r^2 = y_i y_i$

$$\eta = \frac{r}{(ct)^{1/2}}$$

$$c_1 = \frac{(1-2\nu)}{\lambda_s + 2\mu}\frac{\theta_0\beta^2}{\rho c_\epsilon} = \frac{\nu_s - \nu}{1-\nu_s}$$

$$c = \frac{k}{\rho c_\epsilon}\frac{\lambda + 2\mu}{\lambda_s + 2\mu}.$$

For the displacement kernel,

$$G_{ij} = \frac{1}{16\pi r}\frac{1}{\mu(1-\nu)}\left[\left(\frac{y_i y_j}{r^2}\right) g_1(\eta) + (\delta_{ij}) g_2(\eta)\right]$$

$$G_{i\theta} = \frac{1}{4\pi}\left(\frac{\beta}{k(\lambda+2\mu)}\right)\left[\left(\frac{y_i}{r}\right) g_3(\eta)\right]$$

$$G_{\theta j} = \frac{1}{8\pi}\left(\frac{\beta}{k(\lambda+2\mu)}\right)\left[\left(\frac{y_j}{r}\right) g_4(\eta)\right]$$

$$G_{\theta\theta} = \frac{1}{4\pi r}\left(\frac{1}{k}\right)[g_5(\eta)]$$

whereas, for the traction kernel,

$$F_{ij} = \frac{1}{8\pi r^2}\frac{1}{1-\nu}\left[-\left(\frac{y_i y_j y_k n_k}{r^3}\right) f_1(\eta) - \left(\frac{\delta_{ij} y_k n_k + y_i n_j}{r}\right) f_2(\eta) + \left(\frac{y_j n_i}{r}\right) f_3(\eta)\right]$$

$$F_{i\theta} = \frac{1}{2\pi r}\left(\frac{\mu\beta}{k(\lambda+2\mu)}\right)\left[\left(\frac{y_i y_k n_k}{r^2}\right) f_4(\eta) + (n_i) f_5(\eta)\right]$$

$$F_{\theta j} = \frac{1}{8\pi r}\left(\frac{\beta}{\lambda+2\mu}\right)\left[\left(\frac{y_j y_k n_k}{r^2}\right) f_6(\eta) - (n_j) f_7(\eta)\right]$$

$$F_{\theta\theta} = \frac{1}{4\pi r^2}\left[\left(\frac{y_k n_k}{r}\right) f_8(\eta)\right].$$

In the above,

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-x^2} dx = 1 - \operatorname{erfc}(z)$$

$$h_1(\eta) = \operatorname{erf}\left(\frac{\eta}{2}\right) - \frac{\eta}{\sqrt{\pi}} e^{-\eta^2/4} \qquad h_1' = \frac{\partial h_1}{\partial \eta}$$

$$h_2(\eta) = \frac{6h_1(\eta)}{\eta^2} - \frac{\eta}{\sqrt{\pi}} e^{-\eta^2/4}$$

$$g_1(\eta) = 1 + c_1\{h_1(\eta) - h_2(\eta)\}$$

$$g_2(\eta) = (3-4\nu) - c_1\left\{h_1(\eta) - \frac{h_2(\eta)}{3} + \frac{2\eta}{3\sqrt{\pi}} e^{-\eta^2/4}\right\}$$

$$g_3(\eta) = \frac{h_1(\eta)}{\eta^2 t}$$

$$g_4(\eta) = \operatorname{erfc}\left(\frac{\eta}{2}\right) + \frac{2h_1(\eta)}{\eta^2}$$

$$g_5(\eta) = \operatorname{erfc}\left(\frac{\eta}{2}\right)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$f_1(\eta) = 3 + 3c_1\left\{h_1 - \frac{5h_2}{3}\right\}$$

$$f_2(\eta) = (1-2\nu) - c_1\{h_1 - h_2\}$$

$$f_3(\eta) = (1-2\nu) - c_1\{h_1 + h_2\}$$

$$f_4(\eta) = \frac{h_1'}{\eta t} - \frac{3h_1}{\eta^2 t}$$

$$f_5(\eta) = \frac{h_1(\eta)}{\eta^2 t} - \frac{h_1'}{\eta t}$$

$$f_6(\eta) = \operatorname{erfc}\left(\frac{\eta}{2}\right) + \frac{6h_1(\eta)}{\eta^2}$$

$$f_7(\eta) = \operatorname{erfc}\left(\frac{\eta}{2}\right) + \frac{2h_1(\eta)}{\eta^2}$$

$$f_8(\eta) = 1 - h_1(\eta).$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Fatigue, Fractals, and a Modified Miner's Rule

Eli Altus
Faculty of Mechanical Engineering,
Technion-Israel Institute of Technology,
Haifa, 32000 Israel
Assoc. Mem. ASME

A Mechano Chemical Fatigue Model (MCFM) was recently developed. Two basic features of the model are studied here: (a) the fractal (self-similarity under magnification) property of the damage function, which is predicted by the model and (b) how it leads to a "Modified Miner's Rule" (MMR), where the fatigue response is dependent on the loading order. Based on the simplest (basic) case of the model, the damage function is dependent on a single parameter when a Weibull strength distribution for the microscale elements (chains) is used. the MCFM proposes a rational explanation to a linear fatigue failure envelope for a two level (High-Low, Low-High) stress loading while the damage law is nonlinear. This is in contrast to the regular Miner Rule, where the damage progression is assumed linear. Applying multilevel cyclic loading the model shows a "fading memory" response for a monotonically increasing stress. While for a decreasing stress steps there is an intermediate delay stage for which a constant damage rate is found. The microscopic basis of the MCFM enables a clear and direct physical interpretation of the macro behavior as related to the microscale damage mechanisms.

Introduction and Background

A Mechano-Chemical Fatigue Model (MCFM), which has been developed recently (Altus, 1989a,b) originates from a microscale approach where a continuum solid is described by a large set of "elements" or "chains." The chains have a statistical strength distribution such that when a uniform strain is applied, evenly distributed weak chains are broken.

Each site of a broken chain may lead to a successive breakages of its neighbor not merely by the "classical" stress concentration near a hole or a crack, but also by the tendency of the free edges of the broken chains to react with chains in their nearest vicinity. This tendency increases during the unloading stage, when the edges are forced to move relative to their neighbor, creating more "opportunities" (therefore raising the probability) for new bondings. The new structure causes the neighbor to have a nonuniform strain field in addition to the local weakening of the original bond. Therefore, the bonded neighbor loses some of its strength, which increases the probability for a continuing breakage when the next loading is applied. Accumulation of breakages originating from one site can be considered a "crack," such that in each cycle there are two types of damage mechanisms: expansion of "old" cracks and the appearance of new ones. The effect of the crack size on the damage progression is not considered here and will be examined in a separate study.

Using statistical arguments the macro fatigue response is found, which leads to a direct relationship between the microscale structure and the macro parameters. Major advantages of the model includes the ability to predict the stress-life power-law relationship (S−N curve) and the "knee point" phenomenon (endurance limit) in fatigue.

It should be noted that the idea of a successive breakage of microstructure elements has been used extensively by many (Daniels, 1945; Coleman, 1958; Rosen, 1964; Zweben, 1968; Zhurkov, 1974; Phoenix, 1975; to name a few). However, the application to fatigue mechanisms has been rarely studied, and not by the present interpretation (Gohnem and Provan, 1980).

A unique feature of the MCFM is found when the model takes its simplest form (the "basic" case), i.e., when each crack expands by one chain in each cycle: The damage function is identical for all stress levels when plotted against the relative number of cycles, in spite of its strong nonlinearity.

This property and the stress-life power-law prediction capability gave strong motivation for the present study which progresses in two linked directions: (*a*) to understand the characteristics of the above damage function better, and (*b*) to reveal the response of the basic case to a two-level fatigue loading. As will be seen, the damage function was found to possess a self-similarity property, which is the dominant characteristic of fractals. Using the above is a major key for proving important features of the response under two level loading conditions, i.e., a Modified Miner's Rule (MMR), which has been so far an empirical law without a physical microscale mechanism as a basis (Miner, 1945).

Apart from giving a new insight into the micro fatigue processes, the connection between fractal sets and the behavior of physical systems has been an exciting subject in recent years (Mandelbrot, 1982). There are preliminary results (Louis et al., 1985) that the geometry and shape of the damage can be

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 17, 1989; final revision, Jan. 18, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

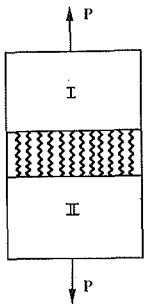


Fig. 1 Schematic representation of a unidirectional parallel elements, or a set of molecular chains

associated with fractals. Any new example showing such a correspondence will add to the understanding of both the mathematical aspects of chaotic behavior and the physical phenomenon.

The Fractal Damage Equation for the MCFM Basic Case

Consider a parallel set of molecular chains loaded unidirectionally by a uniform strain ϵ as shown in Fig. 1. The chains have a strength distribution with a density function $f(\epsilon)$.

Assuming no residual stresses, if the material is loaded from 0 to ϵ, the relative number of broken chains will be:

$$F(\epsilon) = B_1 = \int_0^{\epsilon} f(\epsilon')d\epsilon', \tag{1}$$

where B_n is the relative number of new breakages during cycle n.

The broken edges are unstable and tend to react with their unbroken neighbors. This tendency depends on the specific adjacent structure and the external displacements of the fatigue cycle. It is controlled here by a probabilistic function $g:(0-1)$. For $g=1$, a chemical reaction (or bonding) will definitely happen at each site of a broken edge, and for $g=0$ it will never happen.

The reacted edges reduces the capability of the unbroken neighbor to withstand the same load as before, which is for the present analysis equivalent to a local stress concentration factor k.

When the material is reloaded, part of the "contaminated" chains will break before the global stress reaches the previous one. To sustain the same load, the strain will have to increase beyond the value for the first loading, which will force some of the "nonedge" weak chains to break. This process continues until total separation failure occurs.

The analysis in the following is concentrated on the basic case where the probability for a chemical reaction is 100 percent ($g=1$), and the stress concentration factor k is infinite, meaning that each broken edge will cause another breakage at every cycle. It may simulate materials which have a finite fatigue life for any stress level, since any new cycle involves new breaks. The macro behavior will show no endurance limit. The maximum outer stress in the first loading, which will be kept equal for all cycles is:

$$\sigma_0 = E\epsilon_1[1 - F(\epsilon_1)], \tag{2}$$

where E is the initial elastic modulus when all chains unbroken, or, alternatively, the stiffness of a single chains. The subscript stands for the cycle number. In the second cycle, $2B_1$ chains will be broken when the strain reaches ϵ_1 again. B_1 "old", from the first cycle, and B_1 "new" neighbors. The newly broken chains had a strength distribution density which is identical to all chains which survived the first cycle, since their position is defined arbitrarily prior to any loading. Therefore, their relative strength distribution must be proportional to f. The number of chains survived as a function of their strength is:

$$f(\epsilon < \epsilon_1) = 0,\ f(\epsilon > \epsilon_1) = (1 - \alpha_1)f(\epsilon). \tag{3}$$

α_1 is the proportional coefficient which can be found by equating the number of the second cycle failures to the number of crack sites in the previous cycle:

$$\int_{\epsilon_1}^{\infty} \alpha_1 f(\epsilon)d\epsilon = B_1 = F(\epsilon_1). \tag{4}$$

Therefore,

$$\alpha_1 = F(\epsilon_1)/[(1 - F(\epsilon_1)]. \tag{5}$$

The number of new sites, created from ϵ_1 to ϵ_2 will be

$$\int_{\epsilon_1}^{\epsilon_2} (1 - \alpha_1)f(\epsilon)d\epsilon = (1 - \alpha_1)[F(\epsilon_2) - F(\epsilon_1)]. \tag{6}$$

Equating the outer stress for the first and second cycles and using equations (2), (4), and (6) we get

$$\sigma_0 = E\epsilon_2[1 - 2F(\epsilon_1) - (1 - \alpha_1)(F(\epsilon_2) - F(\epsilon_1))]. \tag{7}$$

When the load is raised at the third time both crack groups, i.e., one and two adjacent broken chains, will create another definite failure at each crack site. Their number is calculated from (1) + (6):

$$(1 - \alpha_1)[F(\epsilon_2) - F(\epsilon_1)] + F(\epsilon_1) = \alpha_1 F(\epsilon_1) + (1 - \alpha_1)F(\epsilon_2). \tag{8}$$

The distribution of strength of the unbroken chains will then be:

$$f(\epsilon < \epsilon_2) = 0;\ f(\epsilon > \epsilon_2) = f(\epsilon)(1 - \alpha_1)(1 - \alpha_2), \tag{9}$$

where α_2 is found similarly by equating the number of the second and third chain failures to the number of crack sites:

$$\int_{\epsilon_2}^{\infty} \alpha_2(1 - \alpha_1)f(\epsilon)d\epsilon = F(\epsilon_1) + (1 - \alpha_1)[F(\epsilon_2) - F(\epsilon_1)]. \tag{10}$$

The number of new sites created during loading from ϵ_2 to ϵ_3 will be:

$$\int_{\epsilon_2}^{\epsilon_3} (1 - \alpha_1)(1 - \alpha_2)f(\epsilon)d\epsilon = (1 - \alpha_1)(1 - \alpha_2)[F(\epsilon_3) - F(\epsilon_2)]. \tag{11}$$

Again, ϵ_3 is found by equating the outer stress to σ_0:

$$\sigma_0 = E\epsilon_3[1 - 3F(\epsilon_1) - 2(1 - \alpha_1)(F(\epsilon_2) - F(\epsilon_1)) - (1 - \alpha_2)(1 - \alpha_1)(F(\epsilon_3) - F(\epsilon_2))]. \tag{12}$$

A definite pattern is revealed by the relations (2), (7), and (12) and the results for any cycle can be found successively. The following relationship is used for the nth cycle:

$$\sigma_0 = E\epsilon_n(1 - D_n), \tag{13}$$

where D_n is the total relative number of broken chains from all cycles which, from the macroscale view can be defined as damage. Total failure occurs when the damage value is 1. The relative number of unbroken chains after n cycles is

$$\eta_n = 1 - D_n, \tag{14}$$

which is also the relative (residual) macro stiffness. Note that equation (14) is compatible with the classical macroscale damage definition proposed by Kachanov (1958).

The current relative modulus is also the displacement ratio

$$\eta_n = \epsilon_0/\epsilon_n,\ \epsilon_0 = \sigma_0/E \tag{15a,b}$$

where ϵ_0 is the strain response for a material with no "weakness," i.e., corresponds to E (or alternatively, for $\eta_0 = 1$).

Equations (2), (7), and (12) can now be written as functions of the residual stiffnesses

$$1 - F(\epsilon_0/\eta_1) = \eta_1, \tag{16a}$$

$$1 - F(\epsilon_0/\eta_2) = \eta_1\eta_2/(2\eta_1 - 1) \tag{16b}$$

$$1 - F(\epsilon_0/\eta_3) = \eta_1\eta_2\eta_3/(2\eta_1 - 1)(2\eta_2 - \eta_1). \tag{16c}$$

By induction, the damage evolution equation for the nth cycle is found as:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$1-F(\epsilon_0/\eta_n)=\prod_{i=1}^{n}[\eta_i/(2\eta_{i-1}-\eta_{i-2})] \tag{16d}$$

where

$$\eta_0=\eta_{-1}=1. \tag{17a}$$

Another form of the above is

$$1-F(\epsilon_0/\eta_n)=[1-F(\epsilon_0/\eta_{n-1})]\eta_n/(2\eta_{n-1}-\eta_{n-2}), \tag{17b}$$

which shows the second-order characteristic of the recursive formula. Also recall that equations (17*a*), (17*b*) come directly from the breakage mechanism of the model. They are general and can be used with any strength distribution function.

A Weibull strength distribution, which has been shown to be valid for many physical processes (Weibull, 1951), is now applied

$$F(\epsilon)=1-\exp(-\alpha\epsilon^{\beta}), \tag{18}$$

where α,β are material constants. By introducing the static failure displacement (Rosen, 1964),

$$\epsilon_s=(\alpha\beta)^{-1/\beta}, \tag{19}$$

the probability function can be normalized to include a single material variable β only:

$$F(\epsilon/\epsilon_s)=1-\exp[-(1/\beta)(\epsilon/\epsilon_s)^{\beta}]. \tag{20}$$

For convenience, and without loss of generality, the normalized strains will be used in the following without special notation. Combining equations (17) and (20), the successive damage equation is

$$\exp[-(1/\beta)(\epsilon_0/\eta_n)^{\beta}]=\prod_{i=1}^{n}[\eta_i/(2\eta_{i-1}-\eta_{i-2})]. \tag{21}$$

Thus, $\eta_n(\sigma_0,\beta;n)$ is a discrete (n integer) damage function, with one microstructure variable β. It was shown elsewhere (Altus, 1989a), that equation (21) yields a power-law relation between the number of cycles which is needed to cause a certain amount of damage and the stress level:

$$\sigma_0'/\sigma_0=(n'/n)^{-1/\beta};\ (\eta_{n'}'=\eta_n). \tag{22}$$

It means that the same damage is expected for the same normalized number of cycles independent of the stress level. When $\eta_n=0$, $n=N$ (the total life) and equation (22) gives the stress-life relationship (The S–N curve), known as Basquin's law (Ashby and Jones, 1985).

$$\sigma_0=CN^{-1/\beta}, \tag{23}$$

where C is a material constant related to $\beta(1+1/\beta$ for $\beta\gg1)$.

Further insight into the damage process can be achieved through a schematic description of the strength distribution density function after each cycle, as shown in Fig. 2. The first cycle creates damage by breaking all chains weaker than a certain strain level ϵ_1, similar to the case of a static monotonic loading (area A_1 in the figure). All A_1 broken chain sites will expand in the second cycle breaking additional chains which have a random strength (area B_2). Therefore, $A_1=B_2$ and A_2 chains are broken due to the higher strain ϵ_2. In the third cycle, all cracks (A_1+A_2) will grow, so that $B_3=A_1+A_2$. Continuing this process, the equation for the nth cycle is:

$$B_n=\sum_{1}^{n-1}A_i. \tag{24}$$

Therefore,

$$\sum_{2}^{n}B_i=A_1(n-1)+A_2(n-2)+\ldots..A_{n-1}. \tag{25}$$

The damage can be expressed as

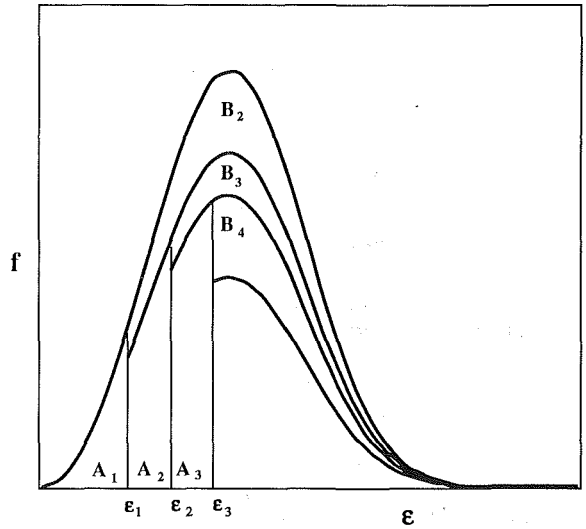


Fig. 2 Representation of the damage evolution through the statistical strength distribution of the chains

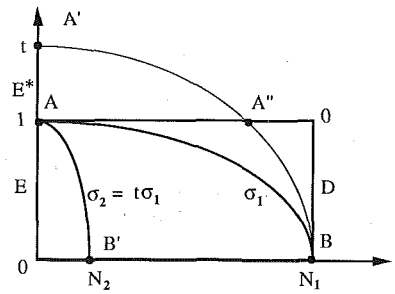


Fig. 3 Damage-life curves showing the self-similarity. Curve *A"B* is an horizontal translation of curve *AB'*, which is a contraction of curve *AB*.

$$D_n=\sum_{i=1}^{n}A_i+\sum_{i=2}^{n}B_i=\sum_{i=3}^{n}(n+1-i)A_i. \tag{26}$$

An important approximation from equations (25) and (26) comes for large enough n (say, greater than 100, which is the common case):

$$D_n\cong\sum_{i=2}^{D}B_i. \tag{27}$$

Notice that the original exponential shape of the strength distribution is unchanged during the cycles.

A unique feature of the damage equation (21) is found by examining the substitution

$$\eta_i=t\eta_i^*. \tag{28}$$

Equation (21) is then written in the form

$$\exp[-1/\beta(\epsilon_0^*/\eta_n^*)^{\beta}]=\eta_n^*t^* \tag{29}$$

where

$$\epsilon_0^*=t\epsilon_0,\quad t^*=t\prod_{i=1}^{n-1}[\eta_i/(2\eta_i-\eta_{i-1})]. \tag{30}$$

It is seen that after a certain number of cycles (i.e., $t^*=1$), equation (29) is exactly the same as the original one (equation (21)), taken for a different stress level ϵ_0^*. The consequence of this result using equation (22) is demonstrated graphically in Fig. 3. Two damage curves for stress levels σ_1 and $\sigma_2=t\sigma_1$ are shown schematically for the same parameter β (lines AB and AB'). From equation (22), line AB' is a horizontal contraction of line AB. The curve $A'B$ is, by definition, a vertical expansion

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of AB (factor of t). However, from equation (30) the curve $A'B$ is identical to AB'!

The above can therefore be summarized as follows: Factorize the original curve vertically and get $A'B$, which is equal to AB'. Then expand horizontally and get the original curve AB. This is a self-similar property under a two-dimensional nonisotropic magnification (contraction). From equation (22), it is characterized by the equation,

$$\eta(n) = t\eta(t^{-\beta}n + m) \tag{31}$$

where t is the expansion factor, β governs the ratio between the horizontal and vertical magnification, and m controls the horizontal translation. m is a function of β and both σ_1 and σ_2.

Fractal properties related directly to the geometry of many physical phenomenons have been discovered in different fields (Mandelbrot, 1982; Cvitanovitz, 1984; Adler, 1986; as examples only). Here, however, the self-similarity is somewhat unique in the sense that it corresponds to a more abstract function which describes the damage evolution of the material.

The Modified Miner Rule (MMR)

The self-similar property above helps in establishing the connection between the damage equation and the MMR for the basic case of the MCFM. The basic problem is the following: Let the loading sequence be defined by a lower cyclic stress level $\sigma_L^{(1)}$ during the first $n_L^{(1)}$ cycles, and then continue with a higher stress $\sigma_H^{(2)}$ till failure. If the number of cycles to failure at each stress level separately (N_H, N_L) is known, how many additional cycles $n_H^{(2)}$ will be left before total failure (the superscript denotes the loading order) occurs?

Recall from equation (22) that a unique normalized damage curve (D versus n/N) is produced for all stress levels, so that

$$D_n(n = n_L^{(1)}) = D_n(n = n_H^{(1)}) \text{ if } n_L^{(1)}/N_L = n_H^{(1)}/N_H. \tag{32}$$

From equations (27) and (32) it is seen that when the stress level is raised from $\sigma_L^{(1)}$ to $\sigma_H^{(2)}$, the amplitude of the strength distribution function of the remaining chains will be the same as if the material was loaded an equivalent number of cycles at $\sigma_H^{(1)}$. Therefore, the state just before the first cycle at $\sigma_H^{(2)}$ is identical to $\sigma_H^{(1)}$ at an equivalent number of cycles, except a difference in ϵ. To prove that this difference disappears after one cycle at $\sigma_H^{(2)}$, we notice that the basic first-cycle equation (16a) will not change if $f(\epsilon)$ includes a cutoff at some level, say, ϵ_c, i.e.,

$$f(\epsilon < \epsilon_c) = 0, \tag{33}$$

since, by definition,

$$1 - F(\epsilon_1) = \int_{\epsilon_1}^{\infty} f(\epsilon)d\epsilon. \tag{34}$$

The expression (34) will remain the same as long as $\epsilon_1 < \epsilon_c$, which corresponds to the case studied here ($L-H$ order). The basic equation at any cycle n remains the same (equation (17)), and therefore the above is true for any stage, i.e., as in equation (34) with ϵ_n.

The above conclusion can now be stated more generally. For a certain damage D and a maximum ϵ (through all loading history), if a new cycle causes $\epsilon_{\max}$ to increase, the residual life will depend solely on D, as if the previous cycles were performed at the new stress level from $D=0$ to the present D. It applies to all cases where the stress levels are monotonically increasing. For a two-stress level case ($L-H$), it leads simply to Miner's rule:

$$n_L^{(1)}/N_L + n_H^{(2)}/N_H = 1. \tag{35}$$

However, while in Miner's equation the order of applying the stress level is not important, here it is:

Consider the ($H-L$) case, where the higher stress level is applied first. The displacements, after increasing monotonically n_H times, will drop to a lower value (say, ϵ_1 in Fig. 2). From the figure it is clear that as long as the displacement will not reach the maximum again, no new cracks are created, only old ones are growing. Therefore, the damage rate will be controlled by the unchanged number of cracks, i.e., a constant value.

Moreover, when the maximum displacement is reached again, the damage progresses "without a memory", the same as in the ($L-H$) case.

All stages can be graphically explained as in Fig. 4. The two damage evolution lines are curves $A-N_L$ and $A-N_H$, corresponding to the stress levels σ_L and σ_H, respectively. Starting from point A, the stress level drops at B, continues with the

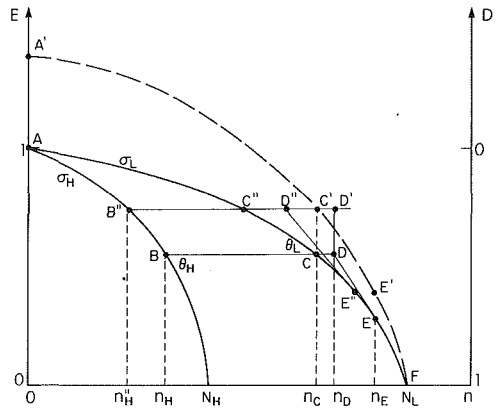


Fig. 4 Schematic damage progression graph for a two level ($H-L$) cyclic loading

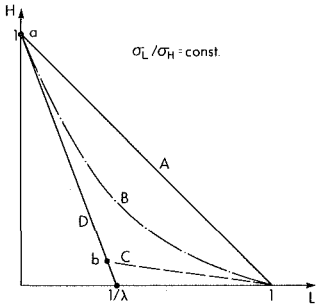


Fig. 5 Schematic representation of a high-low two level fatigue by different theories: (A) Miner (1945), (B) Hashin (1980), (C) Mansor et al. (1986), (D) MMR (present), based on the MCFM

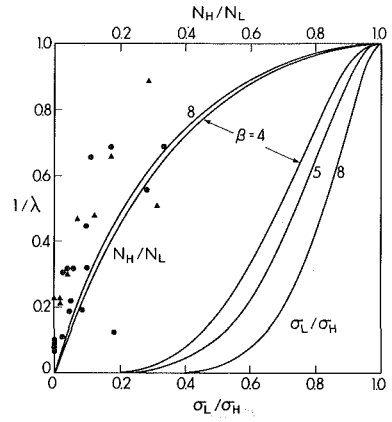


Fig. 6 Variation of $1/\lambda$ as a function of the stress to life ratios for different β. Experimental evidence showing the same trend is taken from Manson et al. (1986). ▲ maraging steel, ● SAE 4130 steel.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

same damage rate, θ_H, until reaching the maximum displacement, point (E), from which the "memory fades away" and the residual life is performed on the original fatigue line (EF). The MMR equation is:

$$n_H^{(1)}/N_H + \lambda n_L^{(2)}/N_L = 1 \tag{36}$$

where λ is dependent on the ratio between the stress levels only. To prove equation (36), we show that (see Fig. 4)

$$R = (n_D - n_C)/(N_L - n_C) = \text{const.} \tag{37}$$

for any number of first-stress level cycles $n_H^{(1)}$.

Let $D(\sigma_L)$ be expanded vertically by an arbitrary factor, t(curve $A'N_L$). Due to the proportional expansion, the ratio between the slopes of the tangent at E and C is the same as the ratio between the slopes at E' and C', which is identical to θ_H/θ_L. As shown above, due to the self-similarity, for any t there is an horizontal factor s (and a translation which does not enter to the calculation here), which brings $C'N_L$ back to CN_L. Again, the slope ratios between the tangents at E'' and C'' is equal to the ratio between the slopes at E' and C', which is identical to the ratio between the tangents of any two points on the damage curves having the same D, like B'' and C''.

It is therefore clear, that the ratio (37) for line CD, is equal to the one for $C''D''$. Finally, since t is arbitrary, R will be constant for any damage level.

The length of the line CD is exactly the number of "lost" cycles, compared to the nonmodified Miner's rule (35). Thus, equation (36) is proved with $R = \lambda$. Moreover, by the same arguments, the above leads to a constant ratios among the lengths $N_L - n_E, n_E - n_D$ and $n_D - n_C$ once N_L/N_H and β is given.

Figure 5 shows three "macro" fatigue theories compared schematically to the MMR for a $H-L$ case (Miner, 1945; Hashin and Laird, 1980; Manson and Halford, 1985). Although the MMR presented here seems to be too simple for a real case, experimental data supply enough evidence for assuming a straight line characteristic in region ab in Fig. 5, especially towards point a. Manson (1980, 1985) even expanded this feature to develop a "double linear damage law."

Since the MMR includes an "internal microstructure," while the above theories do not, it is worthwhile to examine the extent of its validity and the relationship between the slope of the straight line and the microstructure parameters.

The Characteristics of λ

The fractal properties of the damage function leads to some unique features of the MMR which are implicitly embedded in the parameter λ.

From Fig. 4 and equations (13), (14)

$$\sigma_H = E(1 - D_B)\epsilon_B = E\eta_B\epsilon_B. \tag{38}$$

Similarly,

$$\sigma_L = E\eta_E\epsilon_E. \tag{39}$$

From equations (37), (38), and (22) and the arguments in the previous section,

$$\sigma_H/\sigma_L = \eta_B/\eta_E = \eta_D/\eta_E = (N_H/N_L)^{-1/\beta}. \tag{40}$$

The following ratios are defined for convenience:

$$\rho_1 = (N_L - n_E)/(N_L - n_C), \tag{41a}$$

$$\rho_2 = (n_E - n_D)/(N_L - n_C), \tag{41b}$$

$$\rho_3 = (n_D - n_C)/(N_L - n_C) = R, \tag{41c}$$

where $N_I - n_J$, etc., are lengths of the corresponding line segments in Fig. 4. From the self-similarity of the damage function D, and equation (40), we get:

$$\rho_1 = (\eta_D/\eta_E)^\beta = N_{L/H} \tag{42}$$

where $N_{L/H}$ is written instead of N_L/N_H for brevity. Notice that the ratio ρ_1 is dependent on N_H and N_L only, and is free of the only characteristic parameter of the material, β (excluding ϵ_s, which is a scaling factor). It is therefore a *universal property* of the basic MCFM. ρ_2 is calculated next by using the fractal property again. Consider a point x on the damage line AN_L shown in Fig. 4. Then,

$$(N_L - n_x)/(N_L - n_C) = (\eta_x/\eta_E)^\beta. \tag{43}$$

When the point x is sufficiently close to point C, we can define:

$$dN = (N_L - n_x) - (N_L - n_C),\ dE = \eta_x - \eta_C. \tag{44}$$

Then,

$$d\eta/dN = tg(\theta_L) = N_{H/L}tg(\theta_H). \tag{45}$$

Using equations (42)–(45) we get

$$d\eta/dN = \eta_D/(\beta(N_L - n_C)). \tag{46}$$

However, from Fig. 4,

$$tg(\theta_H) = (\eta_D - \eta_E)/(n_E - n_D). \tag{47}$$

Using equations (45)–(47) we finally get

$$\rho_2 = \beta(1 - N_{H/L}^{1/\beta})N_{H/L}. \tag{48}$$

Then,

$$1/\lambda = 1 - \rho_3 = \rho_1 + \rho_2 = N_{H/L}[1 + \beta(1 - N_{H/L}^{1/\beta})]. \tag{49}$$

A specially interesting property of equation (49) is the limit value as β approaches infinity. Using L'Hopital's rule we get

$$\lim_{\beta \to \infty}(1/\lambda) = N_{H/L}(1 - \text{Ln}(N_{H/L})). \tag{50}$$

Equation (49) is shown in Fig. 6 for different β as a function of $N_{H/L}$ and σ_H/σ_L. It is seen that λ is practically independent of β for $\beta > 5$, as predicted by equation (50). This range covers the values for most materials (Ashby and Jones, 1985). Therefore, λ can be considered a *universal value of the model, too.*

To check the validity of the above, experimental data from Manson and Halford (1986) were reduced by measuring the slopes of the lines (ab in Fig. 5) for each experiment. The relatively large amount of data points limits the possibility of receiving "handmade" erroneous trends.

As seen in Fig. 6, the general characteristic of the data fits the theoretical prediction.

Residual Strength

An important parameter of the fatigue resistance is the static residual strength of the material during cyclic loading. From the property in equation (27) and the "fading memory" effect for the $L-H$ case, it can be seen immediately that the residual strength will be linearly dependent on the residual stiffness, i.e.,

$$\sigma_R = E\eta_n\epsilon_s. \tag{51}$$

The similar behavior of $\sigma_R(n)$ and $\eta_n(n)$ was found experimentally for many materials, including laminated composites (Daniel and Charewich 1986). The basic MCFM will predict the above simple relationship for all cases when the strain is monotonically increasing from cycle to cycle and not only for the $L-H$ two-level case.

Conclusions

- The fractal property of the damage function opens a unique way for finding important relationships between fatigue parameters analytically, without having the "mother function" D (which is strongly nonlinear) explicitly. It is a physical example which support the "feeling" (Cvitanovic, 1984; Devaney, 1987) that fractal functions are new tools for understanding many important physical phenomenons.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

- The MMR is a natural expansion of the Miner's rule, in the sense that the "failure envelope" on the $N_H - N_L$ space is a straight line too, both for the $L-H$ and $H-L$ cases.
- For most materials, when $\beta > 5$, the MMR revealed by the basic case of the Mechano Chemical Fatigue Model (MCFM) depends solely on the number of cycles to failure at each separate stress level (N_H, N_L). Some experimental data showed the same trend for the $H-L$ case. However, the validity of the above is limited to a specific region on the fatigue failure space (n/N_L versus n/N_H). For other regions, the simple mechanism of the basic model is insufficient and deserves a separate study.
- The micromechanisms which are inherent in the model enables a direct physical interpretation of the macro behavior in terms of the micro processes. It gives an insight which the "classical" macro theories of fatigue do not have.
- Most fractal functions, which have been found, are related to the physical space (shapes of trees or mountains, the path of a point in a chaotic dynamic case, shapes of turbulent streams, etc.). Here, the self-similarity is found on an abstract space (damage versus life). The ease of getting analytical relationships raises the motivation to search for other nonlinear phenomenons, which can be stated in terms of fractal functions.

Acknowledgment

This work was initiated during the author's sabbatical year with the Institute for Aerospace Studies, University of Toronto. It was partly funded by the Ontario Centre for Material Research, Grant # TP2-325.

References

Adler, P. M., 1986, "Transport Processes in Fractals," *Recent Developments in Structured Media*, (Pitman research notes in mathematics series, #143). D. DeKee and P. N. Kaloni, eds.

Altus, E., 1989a, "Foundation of a Mechano-Chemical Fatigue Theory," UTIAS U. Toronto Report, #330.

Altus, E., 1989, "A Mechano Chemical Model for Fatigue," *J. Mech. of Materials*, accepted for publication.

Ashby, M. F., and Jones, D. R. H., 1985, "Ch. 15 Fatigue Failure," *Engineering Materials, Introduction to Their Properties and Applications*, Pergamon Press, pp. 135–142.

Cvitanovitc, P., ed., 1984, *Universality in Chaos*, Adam Hilger Ltd.

Coleman, B. D., 1958, "Statistics and Time Dependence of Mechanical Breakdown in Fibers," *J. Appl. Phys.*, Vol. 29, pp. 968–983.

Daniels, E. H., 1945, "The Statistical Theory of the Strength of Bundles of Threads," *Proc. R. Soc. London*, Vol. A183, pp. 405–435.

Daniel, I. M., and Charewitcz, A., 1986, "Fatigue Damage Mechanisms and Residual Properties of Graphite/Epoxy Laminates," *Eng. Fract. Mech.*, Vol. 25, pp. 793–808.

Devaney, R. L., 1987, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley Pub., New York.

Gohnem, H., and Provan, W., 1980, "Micromechanics Theory of Fatigue Crack Initiation and Propagation," *Eng. Fract. Mech.*, Vol. 13, pp. 963–977.

Hashin, Z., and Laird, C., 1980, "Cumulative Damage Under Two Level Cycling Loading: Some Theoretical Predictions and Test Data," *Fatigue of Eng. Mat. and Struct.*, Vol. 2, pp. 345–350.

Kachanov, M. L., 1958, "Time of Fracture Process Under Creep Conditions," *Izv. Akad. Nauk SSR O.T.N.*, Tekh. Nauk, Vol. 8, No. 26.

Louis, E., Guinea, F., and Flores, F., 1985, "The Fractal Nature of Fracture," pp. 177–181, *Fractals in Physics*, Proc. 6th Symp., L. Pietronero and E. Torsati, eds., North-Holland Pub., Amsterdam.

Manson, S. S., and Halford, G. R., 1986, "Re-examination of Cumulative Fatigue Damage Analysis—An Engineering Perspective," *Eng. Fract. Mech.*, Vol. 25, pp. 539–571.

Mandelbrot, B. B., 1982, *The Fractal Geometry of Nature*, Freeman, San Francisco.

Miner, M. A., 1945, "Cumulative Damage in Fatigue," ASME JOURNAL OF APPLIED MECHANICS, Vol. 67, pp. A159–A164.

Phoenix, S. L., 1979, "Asymptotic Distribution for the Time to Failure of a Fiber Bundle," *Adv. Appl. Prob.*, Vol. 11, pp. 153–187.

Rosen, B. W., 1964, "Tensile Failure of Fibrous Composites," *AIAA J.*, Vol. 2, pp. 1985–1991.

Weibull, W., 1951, "A Statistical Distribution Function of Wide Applicability," ASME JOURNAL OF APPLIED MECHANICS, Vol. 18, pp. 293–297.

Zweben, C., 1968, "Tensile Failure of Fibrous Composites," *AIAA J.*, Vol. 6, pp. 2325–2331.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

S. She
Graduate Student.

J. D. Landes
Professor.

J. A. M. Boulet
Assoc. Professor.

J. E. Stoneking
Professor,
Mem. ASME.

Department of Engineering Science and Mechanics,
University of Tennessee,
Knoxville, TN 37996

Statistical Theory for Predicting the Failure of Brittle Materials

Statistical models for predicting failure probability of brittle materials are investigated. A formula is derived from a physical consideration for the fracture of microcracks in materials based on the general forms of a fracture criterion and a statistical distribution function incorporating the weakest link principle. The relationships of this model and other statistical models in the literature are discussed; they were found to be equivalent for isotropic materials in which microcracks are randomly distributed in all directions. The statistical model is also used in a failure analysis of the round-notch four-point bending specimen made of an AISI 1008 steel. The grain boundary carbide particles are considered to be microcracks in the plastic zone near the notch tip. The distribution function in the statistical theory is derived from the density and size distribution of carbide particles in the steel. The statistical theory for a triaxial stress state is used to predict the failure probability for any given load on the specimen. The failure loads (loads corresponding to 50 percent of failure probability) are calculated for the specimen at different temperatures. The results are compared with experimental data; good agreement is obtained.

1 Introduction

Brittle materials, such as glass, ceramics, and high strength steel at low temperature, exhibit a substantial dispersion in the fracture strength. Two identical structures made of the same material and under the same loading configuration may have considerably different fracture strengths. This is because the failure of brittle materials is caused by the fracture of microscopic cracks which exist in brittle materials.

The effects of microcracks on macroscopic properties of brittle materials can only be described by statistical theory since the sizes of microcracks are very small and the number of microcracks is very large. Thus, for a loaded structure made of a brittle material, the statistical theory can predict the failure probability under the given load. However, it cannot determine a single failure strength of the structure, because the physical basis for failure of brittle materials is the fracture of the microcracks in the materials.

The objective of this work is to further understand the fracture properties of brittle materials by investigating the statistical models for predicting the failure of these materials. The first step is to investigate the physical basis of the statistical theory for predicting the failure of brittle materials; then the methods for applying the statistical theory for any structure and materials are investigated.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the ASME Winter Annual Meeting, Atlanta, Ga., Dec. 1-6, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 6, 1989; final revision, June 20, 1990.

Paper No. 91-WA/APM-1.

The distribution of microcracks in brittle materials is statistical in nature and has a direct influence on the mechanical properties of the materials. Microcracks in brittle materials can be distributed in the volume or along the surface of the material (Batdorf and Chang, 1979). They can be randomly oriented or have preferred directions. Fracture of microcracks is mostly caused by a normal tensile stress, but it can also be caused by the remote shear stress around the microcracks. The normal compressive stress which can cause the contact of two microcrack surfaces will also effect the fracture of microcracks. The fracture criteria of microcracks are material dependent.

In the current statistical models (Weibull, 1939; Batdorf and Crose, 1974; Evan, 1978; Alpa, 1984), structural failures of brittle materials are associated with fractures of microcracks according to the weakest link principle. The weakest link principle states that the entire structure fails when a single microcrack in the material fails. Application of the weakest link principle to a structure made of brittle material implies that there is no stable microcrack growth in the structure and that any fracture of a microcrack will lead to catastrophic failure of the structure.

Several statistical models have been proposed and applied in predicting the failure of brittle materials. Each one has some success in particular applications, but the limitations and the relations between these models are not clear and have not been investigated extensively.

The Weibull (1939) model was derived from a uniaxial stress state. It is generally used together with the independent action theory in prediction of fracture of brittle materials under multiaxial stress states (Rufin, 1984). The Weibull theory emphasizes the statistical character of the strength of a material. Two statistical distribution functions were proposed by Weibull

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

which used normal tensile stress as a criterion for fracture of microcracks to evaluate the reliability of a structure subjected to tensile stress. The independent action theory assumes that the three principal stresses act independently to affect the failure probability of the loaded body.

Since the 1970s, much research has been done on this problem in relation to the applications of high-strength materials, and some new results were obtained (see Boulet, 1986). Batdorf and Crose (1974) developed a statistical model for a multiaxial stress state from physical consideration of the fracture of microcracks. It was also proved that the Weibull/independent action theory gives nonconservative results (Batdorf, 1977) and the Batdorf and Crose model is more accurate and more realistic than the Weibull/independent action theory (Ruffin et al., 1984). Evans (1978) conducted a statistical analysis by using the strain energy release rate fracture criterion for penny-shaped cracks in triaxial and biaxial stress states. Alpa (1984) proposed a theory in which a general function with several material-dependent parameters is used as a fracture criterion valid for both tension and compression stress states. However, models which are derived from physical consideration of the fracture of microcracks have not been applied to an arbitrary microcrack distribution (Batdorf and Crose, 1974). On the other hand, models which can be applied to an arbitrary microcrack distribution were not derived from a physical consideration of fracture of microcracks (Evans, 1978). In Section 2, these models are reviewed and their relations are investigated. A general formula for the failure probability under a triaxial stress state and for an arbitrary distribution of microcracks and an arbitrary fracture criterion is derived.

Application of the statistical theory to predict the failure of a particular material requires that the mechanism of the fracture process be well understood so that a fracture criterion for the microcracks can be determined and the application of the weakest link principle can be justified. The process of cleavage fracture of AISI 1008 steel is mainly caused by the cracking of grain boundary carbide particles (Lin, 1986; Lin, Evan, and Ritchie, 1986 and 1987). The grain boundary carbide particles initially crack when plastic deformation occurs, and then propagate into the surrounding matrix under a certain stress. The cracked carbide particles are considered to be the major microcracks in the steel, and their size distribution can be found using a microscopic technique called "image analysis." The relationship between the diameter of the crack of a carbide particle and the maximum remote tensile stress the crack can sustain without propagating is also found (Smith, 1966). As an example of applying the statistical theory to a particular material, in Section 3, the statistical theory is used to predict the failure load of a notched body made of AISI 1008 steel from its microstructure properties. This prediction is different from the previous research in that the statistical model for the triaxial stress state is used instead of the previously used statistical model for a uniaxial state.

2 Statistical Models of Fracture of Brittle Materials

Literature Survey. Batdorf and Crose (1974) proposed a statistical model for the multiaxial stress state. This model is derived from the consideration of fracture of microcracks existing in brittle materials. The fracture criterion used is maximum tensile stress. With this criterion, the formula for failure probability is

$$P_f = 1 - \exp\left[-v \int_0^\infty \frac{\Omega(\Sigma, \sigma_{cr})}{4\pi} \frac{dN(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} \right] \tag{1}$$

where $N(\sigma_{cr})$ is the number of microcracks per unit volume having a critical tensile stress less than or equal to σ_{cr}; $\Omega(\Sigma, \sigma_{cr})$ is the solid angle in which normal stress exceeds the value of σ_{cr}; Σ represents the stress state; and v is a small volume element. This model was extended and reformulated for an arbitrary fracture criterion (Batdorf and Heinisch, 1978) and for surface distributed microcracks (Batdorf and Chang, 1979).

Evans (1978) conducted failure analysis for the multiaxial stress state by using the critical coplanar strain energy release rate as a fracture criterion for microcracks.

For the penny-shaped crack,

$$\frac{\pi E G_c}{4(1-\nu^2) r_0} = \sigma_n^2 + \frac{4\tau_n^2}{(2-\nu)^2}, \tag{2}$$

where ν is Poisson's ratio, σ_n is the stress normal to the crack, τ_n is the maximum in plane shear stress, G_c is the critical value of the coplanar strain-energy release rate, and r_0 is the radius of the internal penny-shaped microcrack.

In the analysis, the assumption was made that the three-parameter Weibull distribution function is only valid for the equitriaxial stress state. The strength ratio of equitriaxial stress state to equibiaxial stress state and the strength ratio of equibiaxial stress state to uniaxial stress state were evaluated. Good agreement with experimental data was obtained.

Matsuo (1981) derived the general statistical theory for the multiaxial stress state by assuming that Weibull's three-parameter distribution function is only valid for the equitriaxial tensile stress state which was proposed by Evans. His formulae can be used in the case that the distribution of microcracks in the materials has some preferred directions. Both effects of internal and surface cracks are included in the theory.

The General Formula of Failure Probability. Suppose the fracture criterion of microcracks under any stress state is given; then the failure probability can be calculated from a statistical distribution function which characterizes the fracture strength resulting from the existence of microcracks.

In the following discussion, general forms are assumed for both the fracture criterion of the microcracks and the distribution function. The formula for predicting failure probability is derived from these general forms and the weakest link principle. This formula is equivalent to the formula given by Batdorf and Crose if microcracks are randomly distributed in the materials, and it is equivalent to the formula derived by Matsuo for the Evans model if the three-parameter Weibull distribution function is used for equitriaxial stress state.

Consider a unit volume of material in which stresses are considered constants. In general, the fracture criterion of a crack is a function of both the remote normal stress σ_n and the remote shear stress t_n, so that

$$K = K(\sigma_n, t_n) \tag{3}$$

where K represents the fracture criterion. Suppose K_{cr} is the critical value of K; then the microcrack will fail if $K > K_{cr}$. Let $\Psi(\theta, \phi, K_{cr})$ be the distribution function, for which $[(1/4\pi)\Psi(\theta, \phi, K_{cr}) (\sin\theta\, d\theta\, d\phi)]$ is the probability of occurrence of one microcrack which has the normal direction within the solid angle $(\sin\theta\, d\theta\, d\phi)$ and has the critical value of the fracture parameter larger than or equal to K_{cr}. Therefore, failure probability of the surface, which has a normal direction in the solid angle element $d\Omega = \sin\theta\, d\theta\, d\phi$, is

$$P_f(d\Omega) = \frac{1}{4\Pi} \Psi(\theta, \phi, K_{cr}) (\sin\theta\, d\theta\, d\phi). \tag{4}$$

The survival probability of this element is

$$P_s = 1 - P_f(d\Omega) = 1 - \frac{1}{4\pi} \Psi(\theta, \phi, K_{cr}) (\sin\theta\, d\theta\, d\phi). \tag{5}$$

The survival probability of a unit volume is the multiplication of the survival probabilities of the small elements in all the directions. Suppose that Ω_0 represents the entire solid angle from the center of a unit sphere, that is, $\Omega_0 - 4\Pi$. Then,

$$P_s = \prod_{d\Omega \epsilon \Omega_0} [1 - P_f(d\Omega)]. \tag{6}$$

From equation (6), one has

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\prod_{d\Omega\in\Omega_0}\left[1-P_f(d\Omega)\right]$$

$$=\prod_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\left[1-\frac{1}{4\pi}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (7)$$

From the Taylor expansion of function $\exp(-x)$ when x is small,

$$\exp(-x)=1-x+o(x). \quad (8)$$

Letting x be $1/4\pi\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)$ in equation (8), and noting that $(\sin\theta\,d\theta\,d\phi)$ is small, then

$$1-\frac{1}{4\pi}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)$$

$$=\exp\left[-\frac{1}{4\pi}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (9)$$

Substituting equation (9) into equation (7) and then into equation (6), one obtains

$$P_s=\prod_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\exp\left[-\frac{1}{4\pi}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (10)$$

Changing the multiplication in equation (10) into the summation of the exponents, one has

$$P_s=\exp\left[\sum_{\substack{0<\theta<\pi\\0<\phi<2\pi}}-\frac{1}{4\pi}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (11)$$

In equation (11), supposing $d\theta$ and $d\phi$ are infinitesimal, the summation sign can be changed into integration sign, on has

$$P_s=\exp\left[-\frac{1}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\theta,\phi,K_{cf})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (12)$$

The right-hand side of equation (12) is the survival probability for one crack in the unit volume of materials. If there are N_0 microcracks per unit volume of the material, then from the weakest link principle the total survival probability is the product of the survival probability of each microcrack. One has

$$P_s=\exp\left[-\frac{N_0}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]. \quad (13)$$

Using the same argument above, the survival probability for whole structure is

$$P_s=\exp\left\{-\frac{N_0}{4\pi}\int_v\left[\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]dv\right\}. \quad (14)$$

Finally, from equation (14) the failure probability for the whole structure becomes

$$P_f=1-\exp\left\{-\frac{N_0}{4\pi}\int_v\left[\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\times\Psi(\theta,\phi,K_{cr})\ (\sin\theta\,d\theta\,d\phi)\right]dv\right\}. \quad (15)$$

This is the general formula for predicting the failure probability in the structure in which the stresses are known.

Discussion. The Batdorf and Crose model (see equation (1)) was derived physically from a fracture criterion based on normal tensile stress, a distribution function $N(\sigma_{cr})$ and the weakest link principle for isotropic materials in which microcracks are randomly distributed in all directions. The distribution function $N(\sigma_{cr})$ is defined as the density of microcracks having a critical tensile stress less than or equal to σ_{cr}. The general theory developed in Section 2 can also be applied on the same physical basis. Thus, it should give the same formula for calculating the failure probability if it is applied to isotropic materials and if the normal tensile stress is used as a fracture criterion.

In the theory developed here, if a fracture criterion based on maximum tensile stress is used and material is isotropic, then $K_{cr}=\sigma_{cr}$ and the distribution function Ψ is only a function of σ_{cr}, that is, $\Psi\doteq\Psi(\sigma_{cr})$. From both definitions of distribution function N and Ψ, one has

$$N_0\Psi(\sigma_{cr})=N(\sigma_{cr}). \quad (16)$$

Then, one has

$$\frac{N_0}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\sigma_{cr})(\sin\theta\,d\theta\,d\phi)=\frac{1}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}N(\sigma_{cr})\ (\sin\theta\,d\theta\,d\phi). \quad (17)$$

Since

$$N(\sigma_{cr})=\int_0^{\sigma_{cr}}\frac{dN(\sigma_{cr})}{d\sigma_{cr}}\,d\sigma_{cr},$$

equation (17) can be changed into

$$\frac{N_0}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\sigma_{cr})(\sin\theta\,d\theta\,d\phi)=\frac{1}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\int_0^{\sigma_{cr}}\frac{dN(\sigma_{cr})}{d\sigma_{cr}}\,d\sigma_{cr}\ (\sin\theta\,d\theta\,d\phi). \quad (18)$$

One should notice the different meaning of σ_{cr} at the different positions in the right side of equation (18). The σ_{cr} as an upper limit of the integral is a function of θ, ϕ and stress state Σ, that is,

$$\sigma_{cr}=\sigma_{cr}(\theta,\phi,\Sigma). \quad (19)$$

The σ_{cr} inside the integral sign is only an integration variable and does not change with θ and ϕ.

The order of integration in the right side of equation (18) can be exchanged. After the change of the order of integration, the limit of integral for θ and ϕ is a function of σ_{cr} and stress state Σ and is specified by equation (19). Suppose Ω represents the solid angle in which the integration for θ and ϕ is evaluated, and $d\Omega=\sin\theta d\theta d\phi$. Then equation (18) becomes

$$\frac{N_0}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\sigma_{cr})(\sin\theta\,d\theta\,d\phi)=\int_0^{\infty}\int\int_{\Omega(\sigma_{cr},\Sigma)}\frac{dN(\sigma_{cr})}{d\sigma_{cr}}\,d\Omega\,d\sigma_{cr}. \quad (20)$$

In equation (20) since σ_{cr} in the expression $\dfrac{dN(\sigma_{cr})}{d\sigma_{cr}}$ is only an integration variable and is independent of θ and ϕ, one has

$$\int\int_{\Omega(\sigma_{cr},\Sigma)}\frac{dN(\sigma_{cr})}{d\sigma_{cr}}\,d\Omega\,d\sigma_{cr}=\Omega(\sigma_{cr},\Sigma)\frac{dN(\sigma_{cr})}{d\sigma_{cr}}. \quad (21)$$

Substituting equation (21) into equation (20), one obtains

$$\frac{N_0}{4\pi}\int\int_{\substack{0<\theta<\pi\\0<\phi<2\pi}}\Psi(\sigma_{cr})(\sin\theta\,d\theta\,d\phi)=\frac{1}{4\pi}\int_0^{\infty}\Omega(\sigma_{cr},\Sigma)\frac{dN(\sigma_{cr})}{d\sigma_{cr}}\,d\sigma_{cr}. \quad (22)$$

By substituting equation (22) into equation (15) and assum-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ing that the volume v is small such that stresses are constant in v, equation (1) is obtained. Therefore, equation (15) is equivalent to equation (1) if it is applied to isotropic materials and if the maximum tensile stress fracture criterion is used.

The statistical model developed by Evans and Matsuo can be used for unisotropic materials in which microcracks are not randomly distributed and can have some preferred directions, but the theory is not deduced physically from fracture of individual microcracks. The formula from their model for predicting the failure probability is identical to equation (15) if the three-parameter Weibull distribution function is used for the distribution function Ψ in equation (15). Thus, the theory developed in Section 2.2 is equivalent to the Evans model if a three-parameter Weibull distribution function is used. This also proves that the Batdorf and Crose model and the Evans model are actually equivalent for isotropic materials if same fracture criterion and same form of distribution function are used. In general, for isotropic materials, the models should give the same results when identical forms of the distribution function are used and the constants are evaluated on the same material. Failure to recognize this fact has led some investigators to try to make unnecessary and useless comparisons between two different models (Rufin, etc., 1984).

3 Application of Statistical Model for Predicting the Failure Load in a Notched Body

Introduction. The statistical model developed in Section 2 can be used for predicting the failure probability for brittle materials if stresses in the materials are known. In this section, the theory is used to predict the failure load for a four-point bend specimen of AISI 1008 steel with a rounded notch. The problem was investigated by Lin (1986), where a quantitative method was developed to calculate the failure probability for a given load P from the microscopic properties of the materials, and the statistical model for a uniaxial stress state was used. In this study, the statistical model for a triaxial stress state is used for the prediction of the failure load. The advantage of using the statistical model for the triaxial stress state is that the percentage of the carbide particles participating in the fracture process f can be automatically determined by the statistical model for triaxial stress state rather than be approximately estimated as is the case when the statistical model for unixial stress state is used. Thus, the failure load can be completely determined from the microscopic properties of the steel.

The mechanism of the cleavage fracture in the steel is the propagation of the microscopic crack in the materials. In a simple ferrite/carbide steel, the microscopic cracks are initially developed in carbide particles located at the ferrite grain boundary when the plastic deformation occurs in the adjoining ferrite grains. When the stresses in the materials are high enough, the microcracks in the carbide particle will propagate into the adjoining ferrite grains (see Fig. 1). In the four-point bend specimens of AISI 1008 steel, the propagation of one microcrack from carbide particles into the ferrite grains near the notch tip will cause cleavage fracture of the entire specimen. Thus, the weakest link principle can be applied to predict the material behavior. For more information about the statistical investigation of the fracture of steel, see Beremin (1983), Curry and Knott (1978, 1979), Evans (1983), Ritchie et al. (1973), and Wallin et al. (1984).

The density and the size distribution of the carbide particles in the steel can be determined by a microscopic examination. The local fracture stress, perpendicular to the microcrack in a carbide particle, is found to be a function of microcrack diameter, d_0, and other material properties. Therefore, if the stresses in the specimen are known for a given load P, the failure probability can be determined from the distribution of carbide particles in the material.

Since the carbide particles in the steel are considered to become microcracks only when the plastic deformation occurs in the adjoining ferrite grains, only the stresses in the plastic zone need to be known for obtaining the failure probability. In Lin (1986), an approximate method is used to calculate the stresses in the plastic zone at the notch tip for the four-point bending specimen based on Griffiths and Owen's (1971) finite element analysis. In the method, the plastic zone near the notch tip is determined from plasticity analysis at the notch top by using the slip line theory of elastic-perfectly plastic materials, and strains near the notch tip are assumed to be related to the external load P in the same empirical function which is obtained through the finite element analysis of the same specimen. Then, the stresses in the plastic zone can be obtained through the constitutive equation and the equilibrium equation in the plastic zone.

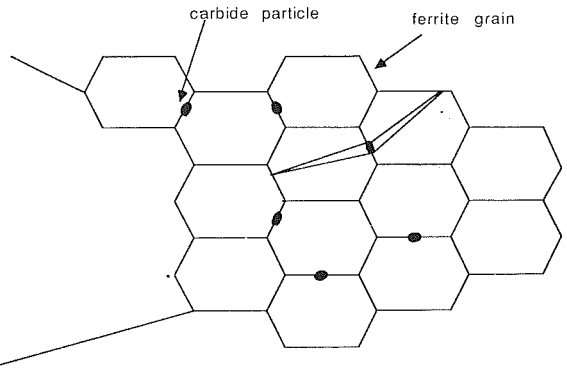


Fig. 1 Schematic illustration of microscopic structure and fracture mechanism of AISI 1008 steel

Mechanism of the Fracture. Smith (1966) developed a theoretical model for obtaining the local fracture stress, α_f, for microcracks which develop from carbide particles. When this model is applied to a penny-shaped microcrack of diameter d_0, the local fracture criterion is found to be

$$\sigma_f = \sqrt{\frac{\pi\gamma E}{(1-\nu^2)d_0}} \tag{23}$$

where γ is the effective fracture surface energy of the surrounding materials, E is the Young's modulus, and ν is Poisson's ratio. This formula has been used in the statistical model to predict the failure load of the notched body of AISI 1008 steel in Lin (1986).

The Stress Field Near the Notch Tip. The size and geometry of the specimen is shown in the Fig. 2. The specimen is a typical notched four-point bend specimen. The distance between two applied forces P is $5W$. The distance between points where displacements in vertical direction are zero is $3W$. The middle section of specimen is subjected to the pure bending of moment PW. The notch depth "a" is one-third of the cross-section. The notch is rounded at the tip and has an opening angle of 45 deg.

The method from Lin (1986) is used here to obtain the stress field near the notch tip of the specimen. In this method, the relationship between applied load and the equivalent strains at the notch as well as the equivalent strain distribution within the plastic zone are obtained from Griffiths and Owen's (1971) finite element analysis; the three principal stresses within the plastic zone near the notch tip are obtained from the equilibrium equations and the constitutive equations of power-hardening materials.

Application of Statistical Theory. When the stress distribution in the plastic zone and microcrack distribution in the material are known, the statistical model can be easily applied.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

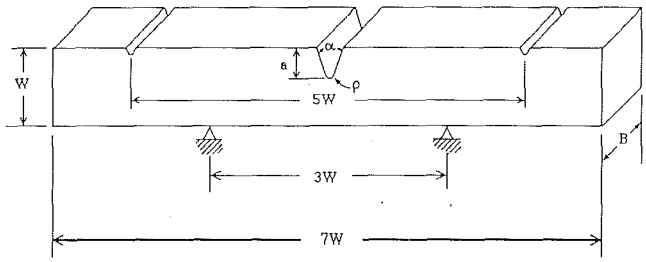


Fig. 2 The geometry of the four-point bending specimen where $a = = 4.23$mm, $B = W = 12.7$mm, $\rho = 0.25$mm, $\alpha = 45$ deg

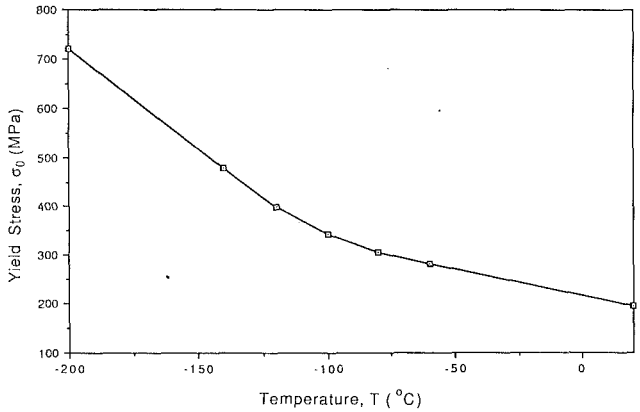


Fig. 3 The change of yield stress, σ_0, with respect to temperature, T

The general formula for obtaining the failure probability P_f is

$$P_f = 1 - \exp\left(-bN_0 \int_{\rho}^{r_p} f(\sigma_1, \sigma_2, \sigma_3) dA\right) \tag{24}$$

where dA is a small element of the area of the plastic zone, b is the characteristic length which is about three times the ferrite grain diameter (see Evan, 1983), N_0 is the number of grain boundary carbide particles per unit volume, and $f(\sigma_1, \sigma_2, \sigma_3)$ is Weibull's distribution function for the multiaxial stress state. In this application, $f(\sigma_1, \sigma_2, \sigma_3)$ has the form

$$f(\sigma_1, \sigma_2, \sigma_3) = -\frac{1}{4\pi} \int\int_{\substack{0<\theta<\pi \\ 0<\phi<2\pi}} \Psi(\sigma_1, \sigma_2, \sigma_3)(\sin\theta \, d\theta \, d\phi) \tag{25}$$

where $\Psi(\sigma_1, \sigma_2, \sigma_3)$ takes the form

$$\Psi(\sigma_1, \sigma_2, \sigma_3) = \begin{cases} \left(\dfrac{\sigma_n - s_u}{s_0}\right)^m & \text{if } \sigma_n > s_u \\ 0 & \text{if } \sigma_n \le s_u \end{cases} \tag{26}$$

where s_u, s_0, m are the three Weibull distribution function parameters. They are determined from the size distribution of carbide particles and equation (23) σ_n is the normal stress in the direction specified by θ and ϕ in the spherical coordinate system of σ_1, σ_2, and σ_3.

$$\sigma_n = \sigma_1 \sin^2\theta\cos^2\phi + \sigma_2 \sin^2\theta \sin^2\phi + \sigma_3 \cos^2\theta \tag{27}$$

If the statistical model for the uniaxial stress state is used as in Lin (1986), then the formula for calculating the failure probability P_f is

$$P_f = 1 - \exp\left(bfN_0 \int_{\pi}^{r_p} G(\sigma_1) dA\right) \tag{28}$$

$$G(\sigma_1) = \begin{cases} \left(\dfrac{\sigma_n - s_u}{s_0}\right)^m & \text{if } \sigma_n > s_u \\ 0 & \text{if } \sigma_n \le s_u. \end{cases} \tag{29}$$

Since the microcracks are randomly distributed in the material, not all the microcracks participate in the fracture process. For instance, if σ_2, σ_3 are less than s_u and the normal direction of a microcrack surface is perpendicular to the direction of the first principal stress σ_1, then this microcrack will not propagate no matter how high σ_1 is. Thus, the parameter f, which is defined as the percentage of the microcracks participating in the fracture process, has to be introduced in the statistical model for uniaxial stress state in order to match the real situation.

The disadvantage of the statistical model for uniaxial stress state is that it does not accurately include the influence of the second and third principal stresses σ_2, σ_3 on the fracture process. Application of the statistical model for the uniaxial stress state requires that the value of f be evaluated from a separate material toughness test and that f be considered a constant independent of stresses in the material. As can be seen from the statistical model for a triaxial stress state, the percentage of the microcracks participating in the fracture process actually depends upon the local stresses; that is, f is not a material property. If in one small material element the three principal stress σ_1, σ_2, σ_3 are same, then all microcracks will equally participate in the fracture process and f should be 1. Thus, the value of f determined from one toughness test can only be applied in the statistical model of the same material under the same loading configuration. This makes prediction of the failure load from the statistical model for uniaxial stress state dependent upon the experimental results of the same material under a similar loading condition. Therefore, the experimental verification of the predicted results from the statistical model for uniaxial stress state in this application cannot be justified logically.

Since the statistical theory for the multiaxial stress state includes the influence of all three principal stresses, there is no need to evaluate the percentage of the microcracks participating in the fracture process f from the results of material toughness tests. Thus, prediction of failure load from the statistical theory of multiaxial stress state depends only upon the microscopic properties of the materials such as size of ferrite grains and distribution of carbide particles in the steel, and upon some other material parameters. The experimental results of the toughness test can be used to verify the predicted failure load from this model.

Results and Discussions. A computer program was developed to calculate the failure probability P_f from a given value of load P and temperature, T, as well as to calculate the value of load P for a given failure probability P_f and temperature. Because the material properties, namely, the yield stress, σ_0, change with temperature, the failure probability will change with temperature for a given load P.

All the materials property parameters needed for the calculation in the program are obtained from Lin (1986). The functional dependency of yield stress, σ_0, upon temperature is plotted in Fig. 3. The size distribution of carbide particles in the steel is shown in Fig. 4. From it, the three Weibull distribution function parameters s_u, s_0 and m are obtained as 1298 MPa, 2838 MPa, and 1.651, respectively. The carbide particle density in AISI 1008 steel is found to be 3.71×10^{-4}/mm^3. The work-hardening exponent is 4.

The load P corresponding to 50 percent, 1 percent, and 99 percent of the failure probabilities are calculated and are plotted in Fig. 5 as a functions of temperature. The results from the statistical model for the multiaxial stress state are quite close to the previous results from the statistical model for the uniaxial stress state with the estimated value of the percentage of microcracks participating in the fracture process f as 5

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

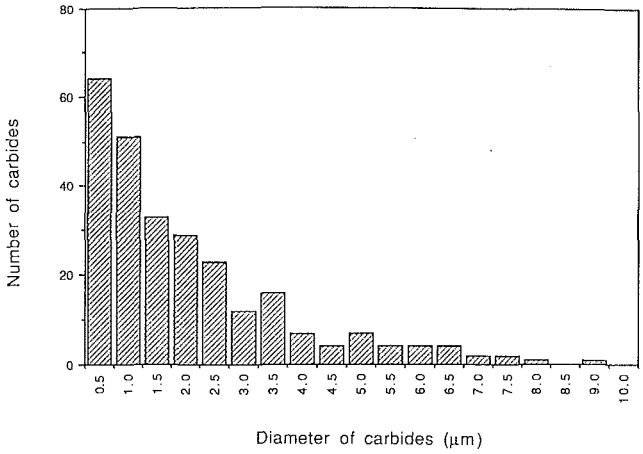


Fig. 4 Size distribution of carbide particles

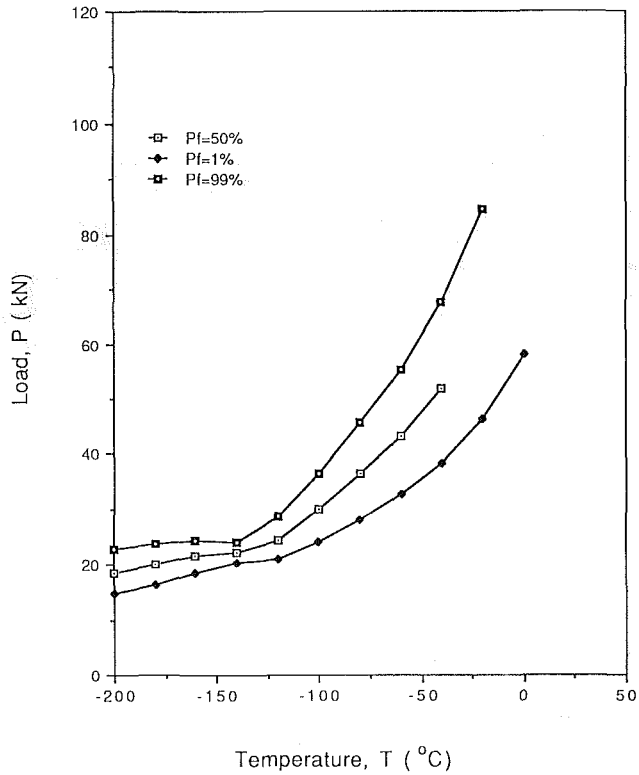


Fig. 5 Predicted loads for failure probabilities, P_f, equal to 1 percent, 50 percent, 99 percent, respectively

percent, and both predicted results are in good agreement with results of the experiments in Lin (1986) at low temperature (see Fig. 6). The agreement of the two models suggests that $f=5$ percent is approximately the correct estimation for the percentage of carbide particles near the notch tip eligible for cracking. The good agreement between the experimental results and the results of the statistical model under a multiaxial stress state at low temperature suggests that this model should be used for the prediction of failure load of brittle materials.

Both the predicted failure load and experimental failure load increase when the temperature increases. This is because the yield stress of the steel decreases as temperature increases; consequently, the stresses near the notch tip will decrease when the yield stress decreases for a given value of load on the specimen. The lower stresses near the notch tip mean that only the microcracks with the largest size are likely to propagate, and thus, fewer microcracks participate in the fracture process

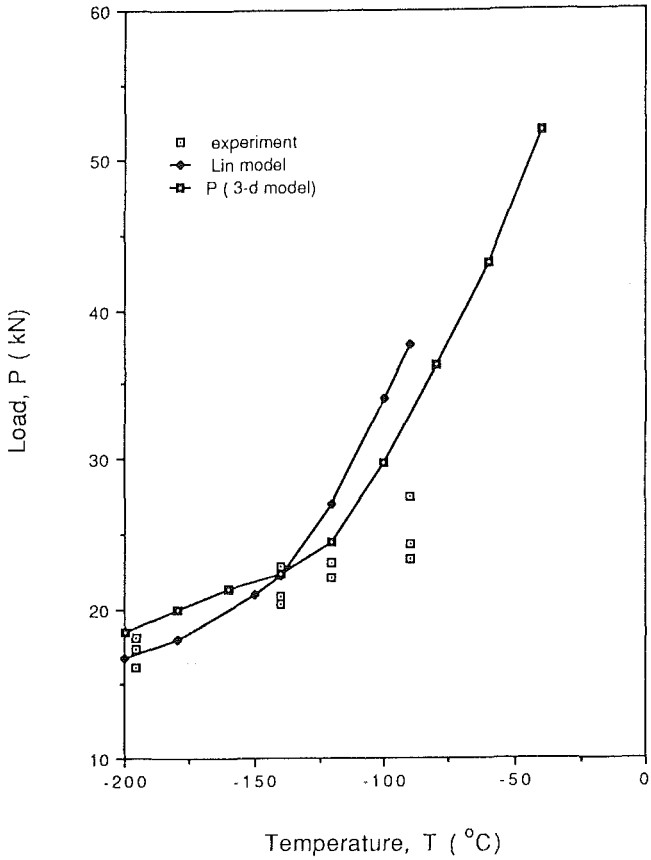


Fig. 6 Comparison of predicted failure loads with measured failure loads

and the specimen is less likely to fail. Therefore, when the temperature increases, the failure load will increase.

4 Conclusions

The statistical theory for brittle materials includes three important elements: the statistical distributions of microcracks in materials, a fracture criterion for microcracks, and the weakest link principle which serves as a failure criterion for the entire loaded body. With arbitrary forms assumed for the distribution function and fracture criterion, the failure probability for a triaxial stress state is obtained mathematically through the weakest link principle. This physically based formula for calculating failure probability is found to be equivalent to the formula given by Batdorf and Crose for isotropic materials and to be equivalent to the formula given by Mastuo for the Evans model if a three-parameter Weibull distribution is used.

Application of the statistical theory for a triaxial stress state to predict failure loads for four-point bending specimens made of AISI 1008 steel also proves to be successful, especially for the failure at low temperature. The theory is applied to an example where the distribution of microcracks is determined directly by analyzing the microstructure of material under the microscope. Using this theory, failure probability can be predicted directly with no additional information. Therefore, the predicted failure load of this model is totally based on the information from the microstructure of the materials and does not have any empirical fitting constants.

References

Alpa, G., 1984, "On a Statistical Approach to Brittle Rupture for Multiaxial States of Stress," *Engineering Fracture Mechanics*, Vol. 19, No. 5, pp. 881–901.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Batdorf, S. B., 1977, "Some Approximate Treatments of Fracture Statistics for Polyaxial Tension," *International Journal of Fracture*, Vol. 13, No. 1, pp. 5–10.

Batdorf, S. B., and Chang, D. J., 1979, "On the Relation Between the Fracture of Volume Distributed and Surface Distributed Cracks," *International Journal of Fracture*, Vol. 15, No. 2, pp. 191–199.

Batdorf, S. B., and Crose, J. G., 1974, "A Statistical Theory for the Fracture of Brittle Structures Subjected to Nonuniform Polyaxial Stresses," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 459–464.

Batdorf, S. B., and Heinisch, H. L., Jr., 1978, "Weakest Link Theory Reformulated or Arbitrary Fracture Criterion," *Journal of The American Ceramic Society*, Vol. 61, No. 7–8, pp. 355–358.

Beremin, F. M., 1983, "A Local Criterion for Cleavage Fracture of a Nuclear Pressure Vessel Steel," *Metallurgical Transactions A*, Vol. 14A, No. 11, pp. 2277–2287.

Boulet, J. A. M., 1986, "An Assessment of the State of Art in Predicting Failure of Ceramics," Final Report, ORNL/SUB/86-57598/1, Oak Ridge National Laboratory, Martin Marietta Energy Systems, Oak Ridge, Tenn.

Curry, D. A., and Knott, J. F., 1978, "Effects of Microstructure on Cleavage Fracture Stress in Steel," *Metal Science*, Vol. 12, pp. 511–514.

Curry, D. A., and Knott, J. F., 1979, "Effects of Microstructure on Cleavage Fracture Toughness of Quenched and Tempered Steel," *Metal Science*, Vol. 13, pp. 341–345.

Evans, A. G., 1978, "A General Approach for the Statistical Analysis of Multiaxial Fracture," *Journal of The American Ceramic Society*, Vol. 61, No. 7–8, pp. 302–308.

Evans, A. G., 1983, "Statistical Aspects of Cleavage Fracture in Steel," *Metallurgical Transactions A*, Vol. 14A, pp. 1349–1355.

Griffiths, J. R., and Owen, D. R. J., 1971, "An Elastic-Plastic Stress Analysis for a Notched Bar in Plane Strain Bending," *Journal of the Mechanics and Physics of Solids*, Vol. 19, pp. 419–413.

Lin, T., 1986, "Stochastic Modeling of Cleavage Fracture Ahead of Rounded Notches and Sharp Cracks," Dissertation, University of California, Berkeley, Calif.

Lin, T., Evans, A. G., and Ritchie, R. O., 1986, "A Statistical Model of Brittle Fracture by Transgranular Cleavage," *Journal of the Mechanics and Physics of Solids*, Vol. 34, No. 5, pp. 477–479.

Lin, T., Evans, A. G., and Ritchie, R. O., 1987, "Statistical Analysis of Cleavage Fracture Ahead of Sharp Cracks and Rounded Notches," *Acta Metall.*, Vol. 34, No. 11, pp. 2205–2216.

Matsuo, Y., 1981, "Statistical Fracture Theory for Multiaxial Stress State Using Weibull's Three-Parameter Function," *Engineering Fracture Mechanics*, Vol. 14, pp. 527–538.

Richie, R. O., Knott, J. F., and Rice, J. R., 1973, "On the Relationship Between Critical Tensile Stress and Fracture Toughness in Mild Steel," *Journal of the Mechanics and Physics of Solids*, Vol. 21, No. 5, pp. 395–410.

Rufin, A. C., Samos, D. R., and Bollard, R. J. H., 1984, "Statistical Failure Prediction Models for Brittle Materials," *AIAA Journal*, Vol. 22, No. 1, pp. 135–140.

Shih, T. T., 1980, "An Evaluation of Probabilistic Approach to Brittle Mechanics," *Engineering Fracture Mechanics*, Vol. 13, pp. 257–271.

Smith, E., 1966, "The Nucleation and Growth of Cleavage Microcracks in Mild Steel," *Proc. Conf. Physical Basis of Yield and Fracture*, Inst. Phys. and Phys. Sci., Oxford, U.K., pp. 34–46.

Wallin, K., Saario, T., and Torronen, K., 1984, "Statistical Model for Carbide Induced Brittle Fracture in Steel," *Metal Science*, Vol. 18, pp. 13–18.

Weibull, W. A., 1939, "A Statistical Theory of the Strength of Materials," *Proc. Roy. Swed. Inst. Eng. Res.*, No. 151.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Proportional Anisotropic Elastic Invariants

A. M. Sadegh
Assoc. Professor,
Mem. ASME

S. C. Cowin
City University of New York
Distinguished Professor,
Fellow ASME

Department of Mechanical Engineering,
City College of the City University
of New York,
New York, NY 10031

There are two proportional invariants for a linear isotropic material, the hydrostatic invariant, and the deviatoric invariant. The former is proportional to the trace of the tensor and the latter is proportional to the trace of the square of the associated deviatoric tensor. The hydrostatic stress and strain and the von Mises stress and strain are directly related to the hydrostatic and deviatoric proportional invariants, respectively, for an isotropic, linear elastic material. For each anisotropic linear elastic material symmetry there are up to six proportional invariants. In this paper we illustrate the six proportional invariants of an orthotropic elastic material using the elastic constants for spruce as the numerical example. The proportional elastic invariants play a role in anisotropic linear elasticity similar to the roles played by the hydrostatic stress and strain and the von Mises stress and strain in isotropic elasticity. They are the unique parameters whose contours represent both the stress and the strain distributions. They also have potential for representing failure or fracture criteria.

Introduction

In the representations of the stress and strain fields that result from the stress analysis of an isotropic elastic object, contour plots of the hydrostatic stress (or dilatational strain) and the von Mises stress (or von Mises strain) are often employed. Also, in the analysis of the flow and fracture of isotropic structural metals the hydrostatic and von Mises stresses have been extremely useful parameters from which considerable physical insight has been obtained. In the case of the plastic deformation of structural metals it was found that, to a first approximation, the yielding was little influenced by a moderate hydrostatic stress. The developments in the plastic theory of structural metals caused considerable emphasis to be placed on the von Mises stress, so named because its importance was first clarified by von Mises (1928). Paul (1968) notes that the von Mises stress has many different names, including the octahedral shear stress.

The hydrostatic stress and strain and the von Mises stress and strain are directly related to the hydrostatic and deviatoric proportional invariants for an isotropic, linear elastic material. The invariants are defined by

$$I_h(\mathbf{T}) = tr\mathbf{T},\ I_h(\mathbf{E}) = tr\mathbf{E}, \tag{1}$$

and

$$I_d(\mathbf{T}) = tr\hat{\mathbf{T}}^2,\ I_d(\mathbf{E}) = tr\hat{\mathbf{E}}^2, \tag{2}$$

where the matrices of tensor components $\mathbf{T}$ ($\mathbf{E}$) and $\hat{\mathbf{T}}$ ($\hat{\mathbf{E}}$) represent the stress (strain) and deviatoric stress (strain), respectively, I stands for invariant, and the h and d subscripts indicate hydrostatic and deviatoric, respectively. The following well-known formulas from isotropic elasticity demonstrate the proportionality of the invariants:

$$I_h(\mathbf{T}) = (3\lambda + 2\mu)I_h(\mathbf{E}),\ I_d(\mathbf{T}) = 4\mu^2 I_d(\mathbf{E}), \tag{3}$$

where λ and μ are the Lamé moduli. The hydrostatic stress, p, and the von Mises stress, σ_{vm}, are related to the proportional invariants by the following formulas:

$$p = \frac{1}{3} I_h(\mathbf{T}) = \frac{1}{3}(T_1 + T_2 + T_3), \tag{4}_1$$

$$\sigma_{vm} = 3\sqrt{\frac{1}{2} I_d(\mathbf{T})} = \sqrt{\frac{3}{2}[(T_1 - T_2)^2 + (T_1 - T_3)^2 + (T_2 - T_3)^2]}, \tag{4}_2$$

where T_1, T_2, and T_3 represent the principal stresses. The hydrostatic strain (or dilatation) and von Mises strain are given by analogous formulas involving the proportional invariants in strain.

We observe that, for an anisotropic elastic material, while it is still possible to form the hydrostatic and von Mises stresses and strains, the major advantages of the formulation are lost because the hydrostatic invariant and deviatoric invariant are

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16–19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 16, 1989; final revision, May 7, 1990.
Paper No. 91-APM-13.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

coupled and no longer proportional. That is to say, equations of the form (3) no longer hold; the hydrostatic stress invariant is no longer proportional to the hydrostatic strain invariant and the deviatoric stress invariant is no longer proportional to the deviatoric strain invariant. This coupling is discussed by Cowin (1990). It should also be noted that the exceptional utility of the von Mises stress in the stress analysis of isotropic metals does not extend to anisotropic materials because the yield or fracture condition for anisotropic materials is not, in general, independent of hydrostatic pressure. There are a great deal of experimental data on natural and manmade anisotropic composite materials and these data demonstrate the influence of hydrostatic stress on the yield and ultimate strength of these materials in contrast to the data on isotropic metals. These data are reviewed by Rowlands (1985). In the special case of bone and wood, which are natural anisotropic composites, some data are given by Cowin (1979).

The early theoretical yield and strength criteria due to von Mises (1928) and Hill (1950) assumed that the yielding or ultimate strength was dependent upon the anisotropy of the material, but independent of the hydrostatic pressure the material was subjected to. The more recent theoretical criteria, including those of Malmeisters (1966) and Tsai and Wu (1971) (see also Wu, 1974), assume that the yielding or strength is dependent upon the anisotropy of the material, but they do not make the assumption of independence of the hydrostatic pressure the material is subjected to. The criteria of this latter type fit the experimental data much more satisfactorily, and they have won relatively universal acceptance in the composite material community (Rowlands, 1985). This body of experimental data and theory clearly demonstrates that the yield or strength criterion of an anisotropic material depends upon the elastic symmetry of the material and that these criteria are not independent of the hydrostatic pressure the material is subjected to. There are intuitive arguments that would lead one to suspect that the strength of an anisotropic material should *not* be independent of the applied hydrostatic pressure. For example, recall that in an anisotropic material a hydrostatic stress state does not produce a state of strain that is a pure dilatation and vice versa. Thus, it is possible that a high state of hydrostatic stress applied to an anisotropic material will produce large shearing strains.

That the elastic anisotropy of a material does significantly influence the subsequent plastic deformation of the material is well known. This observation stems directly from studies of pure crystals in which slip is preferred on certain planes in the crystal.

The purpose of this paper is to demonstrate the significance and illustrate the importance of the proportional invariants of anisotropic linear elasticity. We accomplish these goals, in part, by presenting contour plots of the six proportional invariants of orthotropic elastic symmetry. The stress and strain fields used in the illustration are those around a circular hole in a flat plat subject to a field of uniaxial tension. The orthotropic material constants employed are those of the wood from the spruce tree. We also present contour plots of the hydrostatic stress and strain and the von Mises stress and strain for the same stress and strain fields to illustrate their nonproportionality when the material has orthotropic symmetry. We then compare the contour plots just described with the contour plots of the hydrostatic stress and strain and the von Mises stress and strain for the stress and strain fields around a circular hole in a flat isotropic plate subject to a field of uniaxial tension to illustrate the proportionality when the material has isotropic symmetry.

The concept of the proportional invariants for the various anisotropic elastic symmetries is developed in the next section. In the section that follows, the calculation of the stress and strain fields is described. In the final section there is a discussion of results.

The Proportional Invariants

We begin our discussion of the proportional invariants for anisotropic linear elastic materials by introducing the anisotropic form of Hooke's law. The generalized Hooke's law can be written as

$$T_{ij} = C_{ijkm} E_{km}, \tag{5}$$

where the C_{ijkm} are the components of the elasticity tensor. There are three important symmetry restrictions on the tensor C_{ijkm}. These restrictions require that components with the subscripts *ijkm*, *jikm*, and *kmij* be equal. These restrictions follow from the symmetry of the stress tensor, the symmetry of the strain tensor, and the requirement that no work be produced by the elastic material in a closed loading cycle, respectively. Written as a linear transformation in six dimensions, Hooke's law has the representation, $\mathbf{T} = \mathbf{cE}$, or

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ T_{23} \\ T_{13} \\ T_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ 2E_{23} \\ 2E_{13} \\ 2E_{12} \end{bmatrix}, \tag{6}$$

where the relationship of the components of C_{ijkm} to the components of the matrix $\mathbf{c}$ is easily seen. The inverse of the elasticity tensor $\mathbf{c}$ is the compliance tensor $\mathbf{s}$ where $\mathbf{s} = \mathbf{c}^{-1}$ and $\mathbf{E} = \mathbf{sT}$. Equation (6) can be rewritten in the form

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ \sqrt{2}T_{23} \\ \sqrt{2}T_{13} \\ \sqrt{2}T_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \sqrt{2}c_{14} & \sqrt{2}c_{15} & \sqrt{2}c_{16} \\ c_{12} & c_{22} & c_{23} & \sqrt{2}c_{24} & \sqrt{2}c_{25} & \sqrt{2}c_{26} \\ c_{13} & c_{23} & c_{33} & \sqrt{2}c_{34} & \sqrt{2}c_{35} & \sqrt{2}c_{36} \\ \sqrt{2}c_{14} & \sqrt{2}c_{24} & \sqrt{2}c_{34} & 2c_{44} & 2c_{45} & 2c_{46} \\ \sqrt{2}c_{15} & \sqrt{2}c_{25} & \sqrt{2}c_{35} & 2c_{45} & 2c_{55} & 2c_{56} \\ \sqrt{2}c_{16} & \sqrt{2}c_{26} & \sqrt{2}c_{36} & 2c_{46} & 2c_{56} & 2c_{66} \end{bmatrix} \times \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ \sqrt{2}E_{23} \\ \sqrt{2}E_{13} \\ \sqrt{2}E_{12} \end{bmatrix}, \tag{7}$$

where the six by six matrix contained in this equation is denoted by $\hat{\mathbf{c}}$. The matrix $\hat{\mathbf{c}}$ can be shown to represent the components of a second rank tensor in a six-dimensional space. The proportional invariants are associated with the magnitudes of the eigenvectors of $\hat{\mathbf{c}}$ in the case of strain and with the inverse of this matrix, denoted by $\hat{\mathbf{s}}$, in the case of stress.

In order to give a definition of a proportional invariant we introduce the notation $f(\mathbf{A}, \hat{\mathbf{a}})$ to denote a scalar-valued invariant function of a symmetric second rank tensor in three dimensions, denoted by $\mathbf{A}$, and a symmetric second rank tensor in six dimensions, denoted by $\hat{\mathbf{a}}$. A function $f(\mathbf{A}, \hat{\mathbf{a}})$ is said to be a *proportional invariant* for a particular anisotropic linear elastic material symmetry (e.g., orthotropy, transverse isotropy) if $f(\mathbf{T}, \hat{\mathbf{s}})$ is directly proportional to $f(\mathbf{E}, \hat{\mathbf{c}})$. The constant of proportionality will be the eigenvalues of $\hat{\mathbf{c}}$, or the squares of those eigenvalues. The eigenvalues of $\hat{\mathbf{c}}$ are the inverse of

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

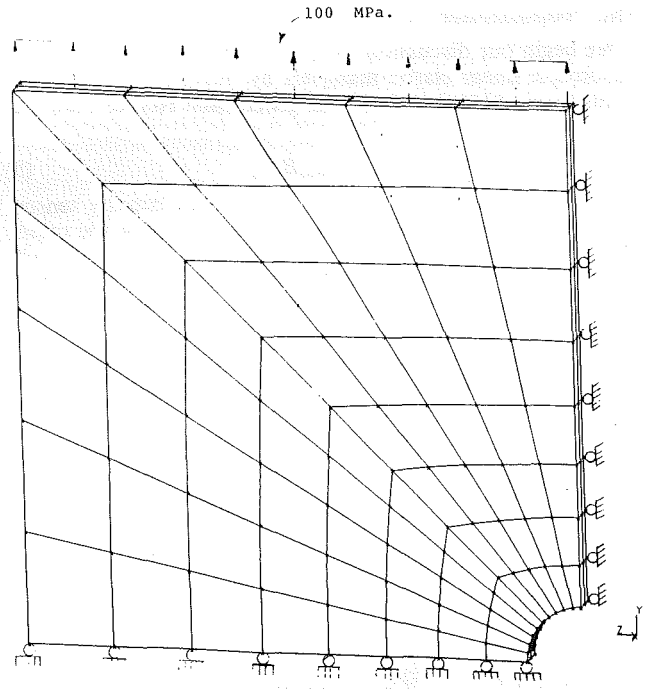


Fig. 1 The finite element model

the eigenvalues of $\hat{\mathbf{s}}$. Equation (3) shows that $I_h(\mathbf{A})$ and $I_d(\mathbf{A})$ are proportional invariants for linear isotropic symmetry and that, in the case of isotropic symmetry, the proportional invariants do not depend on the associated symmetric second rank tensor in six dimensions denoted by $\hat{\mathbf{a}}$ (i.e., the strain proportional invariant does not depend upon $\mathbf{c}$ or $\hat{\mathbf{c}}$, and the stress proportional invariant does not depend upon $\mathbf{s}$ or $\hat{\mathbf{s}}$). In the case of isotropy, the constants of proportionality are $(3\lambda+2\mu)$ and $4\mu^2$, respectively. The proportional invariants for each linear anisotropic elastic symmetry have the same basic properties as the hydrostatic and von Mises stress or strain in the case of isotropic elastic symmetry, but the proportional invariants for the linear anisotropic elastic symmetries may depend on the elastic constants $\mathbf{c}$ or $\mathbf{s}$ while the isotropic proportional invariants are independent of the isotropic elastic constants. The proportional invariants of the anisotropic elastic material symmetries have been identified and calculated for each symmetry by Mehrabadi and Cowin (1990). In the work of Mehrabadi and Cowin, the proportional invariants are called "the invariants of the eigentensors" of the anisotropic elastic material symmetries.

Irreducible sets of the invariants of any second rank tensor $\mathbf{A}$ for each of the anisotropic elastic symmetries are tabulated, for example, by Green and Atkins (1960). An irreducible set of the isotropic invariants of $\mathbf{A}$ are $\mathbf{trA}$, $\mathbf{trA}^2$, and $\mathbf{trA}^3$. Of course, any function of these quantities is also an isotropic invariant of $\mathbf{A}$. The proportional invariants of a particular anisotropic linear elastic material symmetry are particular functions of the invariants in the irreducible sets of invariants listed by Green and Atkins (1960) for the appropriate anisotropic symmetry. In general, there are six proportional invariants for each symmetry because the invariants are associated with the magnitudes of the eigenvectors of $\hat{\mathbf{c}}$ in the case of strain and with the inverse of this matrix in the case of stress. The details of this formulation are given by Mehrabadi and Cowin (1990). In the case of isotropic elastic symmetry, five of the proportional invariants are identical, so there remain only the two distinct invariants given in equation (3).

The proportional invariants of the anisotropic linear elastic material symmetries are two kinds, linear and quadratic. In the case of orthotropic elastic materials there are three linear proportional invariants of a complicated algebraic form and three quadratic proportional invariants of a very simple algebraic form. The linear proportional strain invariants are

$$I_k(\mathbf{E}, \hat{\mathbf{c}}) = K[c_{12}c_{23} + c_{12}c_{13} - c_{12}^2 - c_{23}(c_{11} - \Lambda_k) - c_{13}(c_{22} - \Lambda_k) + (c_{22} - \Lambda_k)(c_{11} - \Lambda_k)], \quad (8)$$

where $k, i, j = 1, 2, 3$, $(k \neq j \neq i)$ and the coefficient $K = K(\Lambda_k, \Lambda_j, \Lambda_i)$ is given by

$$K = K(\Lambda_k, \Lambda_j, \Lambda_i). = \frac{\{N - M\}}{(\Lambda_k - \Lambda_i)(\Lambda_k - \Lambda_j)[c_{12}(c_{13}^2 - c_{23}^2) - c_{13}c_{23}(c_{11} - c_{22})]}, \quad (9)$$

where N and M are given by

$$N = [c_{12}c_{13} - c_{23}(c_{11} - \Lambda_i)][(c_{11} - \Lambda_j)E_{11} + c_{12}E_{22} + c_{13}E_{33}] \quad (10)$$

and

$$M = [c_{12}c_{23} - c_{13}(c_{22} - \Lambda_i)][(c_{22} - \Lambda_j)E_{22} + c_{12}E_{11} + c_{23}E_{33}], \quad (11)$$

and where the Λ_k stands for one of the three eigenvalues of the 3×3 submatrix in the upper left-hand corner of the $\hat{\mathbf{c}}$ representing orthotropic symmetry,

$$\hat{\mathbf{c}} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{22} & c_{23} & 0 & 0 & 0 \\ c_{13} & c_{23} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2c_{66} \end{bmatrix}. \quad (12)$$

As one can see from the formulas (8) to (12), the proportional strain invariants $I_k(\mathbf{E}, \hat{\mathbf{c}})$ given by (8) are linear in the components of the strain tensor $\mathbf{E}$, but rather more complicated functions of the elements of the elasticity tensor $\mathbf{c}$. The linear proportional stress invariants $I_k(\mathbf{T}, \hat{\mathbf{s}})$ are given by similar formulas in which the components of strain $\mathbf{E}$ are replaced by the components of stress $\mathbf{T}$, and the components of $\hat{\mathbf{c}}$ are replaced by the components of $\hat{\mathbf{s}}$.

In general, the quadratic proportional invariants are some subset of the sum of all possible binary products of the components of $\mathbf{A}$; each summand of the sum is multiplied by a scalar coefficient that is an explicit algebraic function of the anisotropic elastic constants of that particular symmetry. In the case of orthotropic elastic materials, the three quadratic proportional invariants are of a very simple algebraic form; they are give by

$$I_4(\mathbf{E}) = E_{23}^2,\ I_5(\mathbf{E}) = E_{13}^2,\ I_6(\mathbf{E}) = E_{12}^2. \quad (13)$$

Similar formulas apply for stress. As was the case with all the isotropic proportional invariants, the quadratic invariants for orthotropic symmetry turn out to be independent of the tensors of elastic constants, $\mathbf{c}$ and $\mathbf{s}$.

The Illustrative Model and Results

In order to illustrate the significance of the proportional invariant concept, we consider a square plate of spruce (20 mm × 20 mm) containing a circular hole of radius 1 mm at its center. The plate is 0.5-mm thick and is subjected to a uniaxial tensile stress of 100 MPa. The axes of the model are chosen such that the stiffer and stronger direction, along the grain, lies in the y-direction in the plane of the plate. The finite element method was employed to determine the stress and strain fields. The finite element model of the problem is shown in Fig. 1, where the thickness is two elements deep. Eight-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

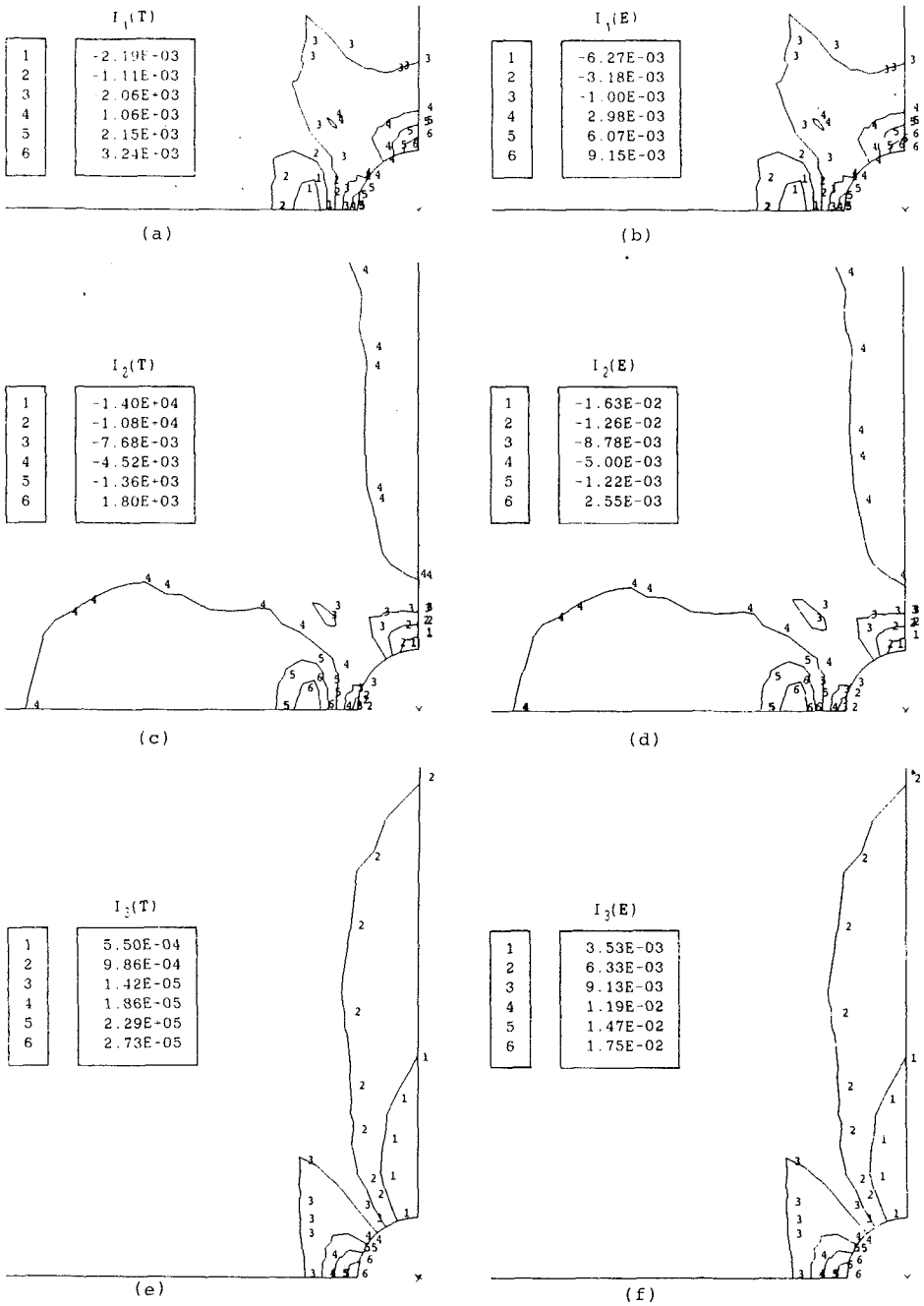


Fig. 2 The three linear proportional stress and strain invariants (the first, second, and third) of spruce: (*a*) first stress invariant, (*b*) first strain invariant, (*c*) second stress invariant, (*d*) second strain invariant, (*e*) third stress invariant, (*f*) third strain invariant

Eight-noded solid brick elements were used for this model. The orthotropic elastic constants for spruce are given by

$$\mathbf{c} = \begin{bmatrix} 0.44 & 0.32 & 0.192 & 0 & 0 & 0 \\ 0.32 & 16.266 & 0.445 & 0 & 0 & 0 \\ 0.192 & 0.445 & 0.78 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.61 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.039 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.76 \end{bmatrix}, \quad (14)$$

in GPa, and the matrix $\hat{\mathbf{c}}$ defined in equation (7) is given by

$$\hat{\mathbf{c}} = \begin{bmatrix} 0.44 & 0.32 & 0.192 & 0 & 0 & 0 \\ 0.32 & 16.266 & 0.445 & 0 & 0 & 0 \\ 0.192 & 0.445 & 0.78 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.22 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.078 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.52 \end{bmatrix}. \quad (15)$$

The eigenvalues of the matrix $\hat{\mathbf{c}}$ are 0.3527, 0.85, 16.31, 1.22, 0.078, and 1.52, also in units of GPa.

The finite element method for orthotropic materials was

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

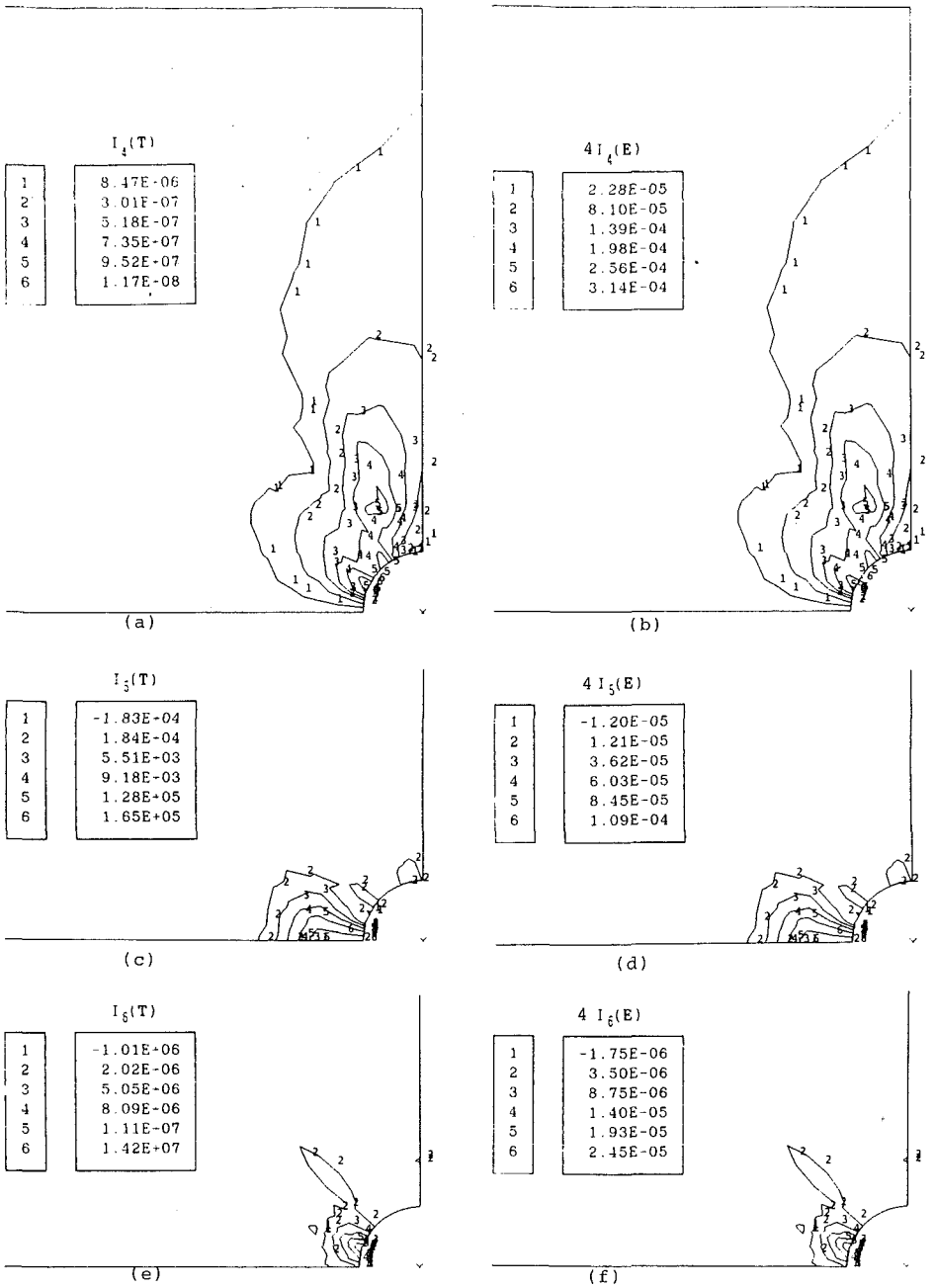


Fig. 3 The three quadratic proportional stress and strain invariants (the fourth, fifth, and sixth) of spruce: (*a*) fourth stress invariant, (*b*) fourth strain invariant, (*c*) fifth stress invariant, (*d*) fifth strain invariant, (*e*) sixth stress invariant, (*f*) sixth strain invariant

employed and the stress and strain tensors of each element were determined. The eigenvalues, the elastic constants, and the strain tensor were substituted into equations (8) and (13) and the strain invariants were calculated. The contour plots of the strain invariants for spruce are given in Figs. 2(*b*), 2(*d*), 2(*f*), 3(*b*), 3(*d*), and 3(*f*). In all the illustrations of this paper, due to the space limitation, only the region of the stress or strain fields where the contour variation of invariants is significant (i.e., where the contour plots are concentrated) is shown. Roughly, the areas in the lower right quarter of the stress and strain fields are shown and the areas at the top left quarter of these fields are omitted. The magnitudes of the proportional stress or strain invariants on each contour are given in the small table in each figure. To calculate the proportional invariants of the stress tensor we replace the strain $\mathbf{E}$ by the stress $\mathbf{T}$ and the elements of $\hat{\mathbf{c}}$ are replaced by the components of $\hat{\mathbf{s}}$. The compliance tensor $\mathbf{s}$ for spruce is given by

$$\mathbf{s} = \begin{bmatrix} 2.564 & -0.034 & -0.613 & 0 & 0 & 0 \\ -0.034 & 0.063 & -0.028 & 0 & 0 & 0 \\ -0.613 & -0.028 & 1.449 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.639 & 0 & 0 \\ 0 & 0 & 0 & 0 & 25.64 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.31 \end{bmatrix}. \quad (16)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

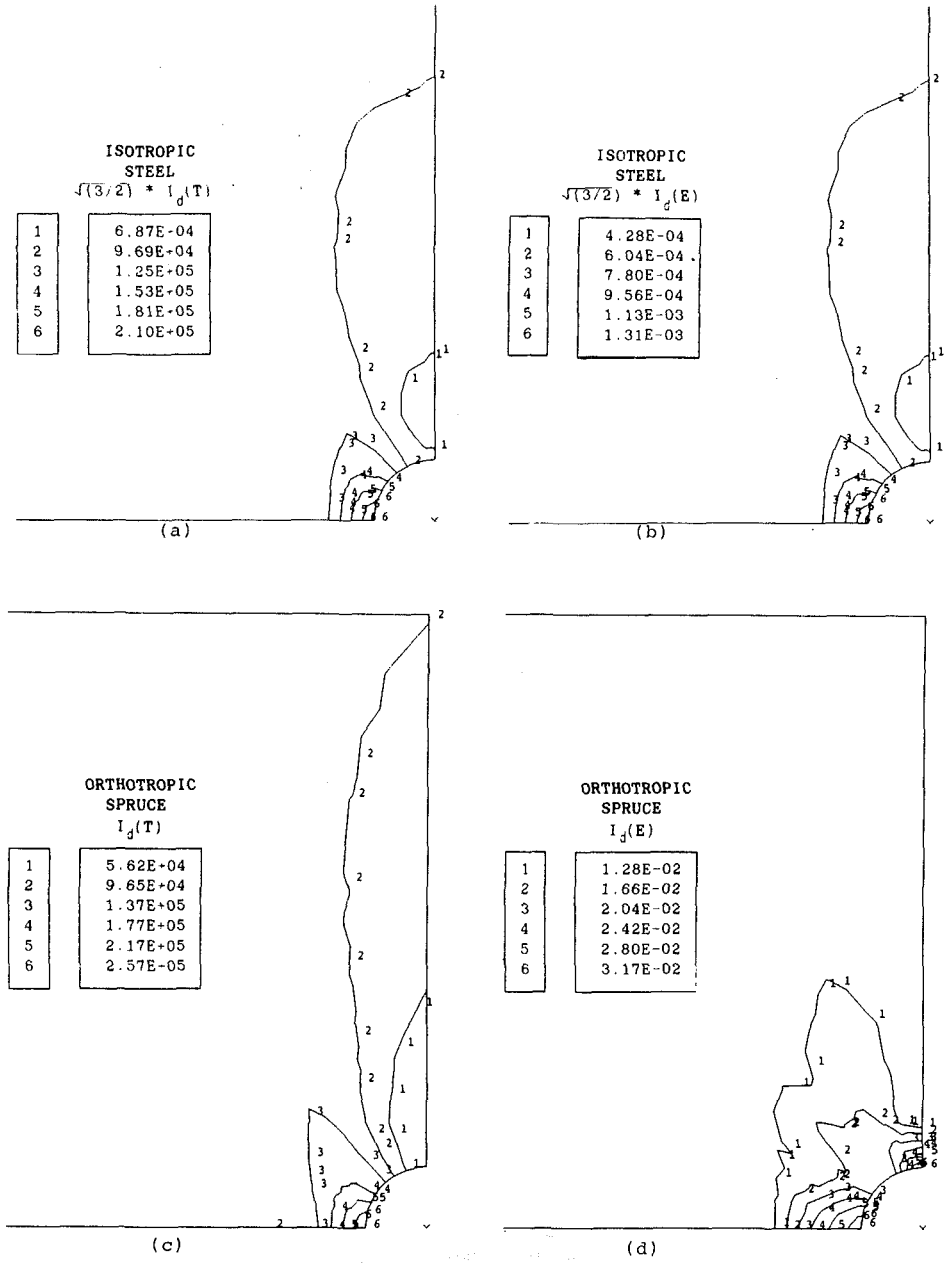


Fig. 4 The von Mises stress and strain for the steel and spruce examples: (*a*) von Mises stress, steel, (*b*) von Mises strain, steel, (*c*) von Mises stress, spruce, (*d*) von Mises strain, spruce

The contour plots of the proportional stress invariants for the spruce are shown in Figs. 2(*a*), 2(*c*), 2(*e*), 3(*a*), 3(*c*), and 3(*e*). The hydrostatic and von Mises stresses were calculated from the formulas (4), and similar formulas were used for the calculation of the hydrostatic and von Mises strains. The contour plots of the von Mises and hydrostatic parts of stresses and strains are shown in Figs. 4(*c*) and 4(*d*), and Figs. 5(*c*) and 5(*d*), respectively.

In order to compare the proportionality of the von Mises and hydrostatic parts of stress and strain in an isotropic material, the computational calculation was repeated for a steel; steel is considered to be isotropic with a Young's modulus of 206.8 GPa and a Poisson's ratio of 0.29. The calculation employed the same geometry and loading as that for the spruce, and the same finite element model was used. The von Mises and hydrostatic parts of stress and strain for steel are shown in Figs. 4(*a*), 4(*b*), 5(*a*), and 5(*b*).

Discussion

The contour plots of the six proportional stress invariants of the orthotropic spruce shown in Figs. 2(*a*), 2(*c*), 2(*e*), 3(*a*), 3(*c*), and 3(*e*) are proportional to the six respective proportional strain invariants of spruce shown in Figs. 2(*b*), 2(*d*), 2(*f*), 3(*b*), 3(*d*), and 3(*f*). Please note that the last three proportional strain invariants of spruce, those shown in Figs. 3(*b*), 3(*d*), and 3(*f*), are multiplied by a factor of four in the figures. The proportionality of the stress and strain invariants can be readily seen by, first, comparing the shape of each contour plot of the stress invariant to that of the strain invariant and, second, by noting that the ratio of the magnitude of the contour of each stress invariant is proportional to the corresponding strain invariant contour. The proportionality constant, that is to say the ratio of the stress invariant to the strain invariant, is equal to the corresponding eigenvalue of $\hat{\mathbf{c}}$ for the linear proportional invariants, the first, second, and third. These numbers are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

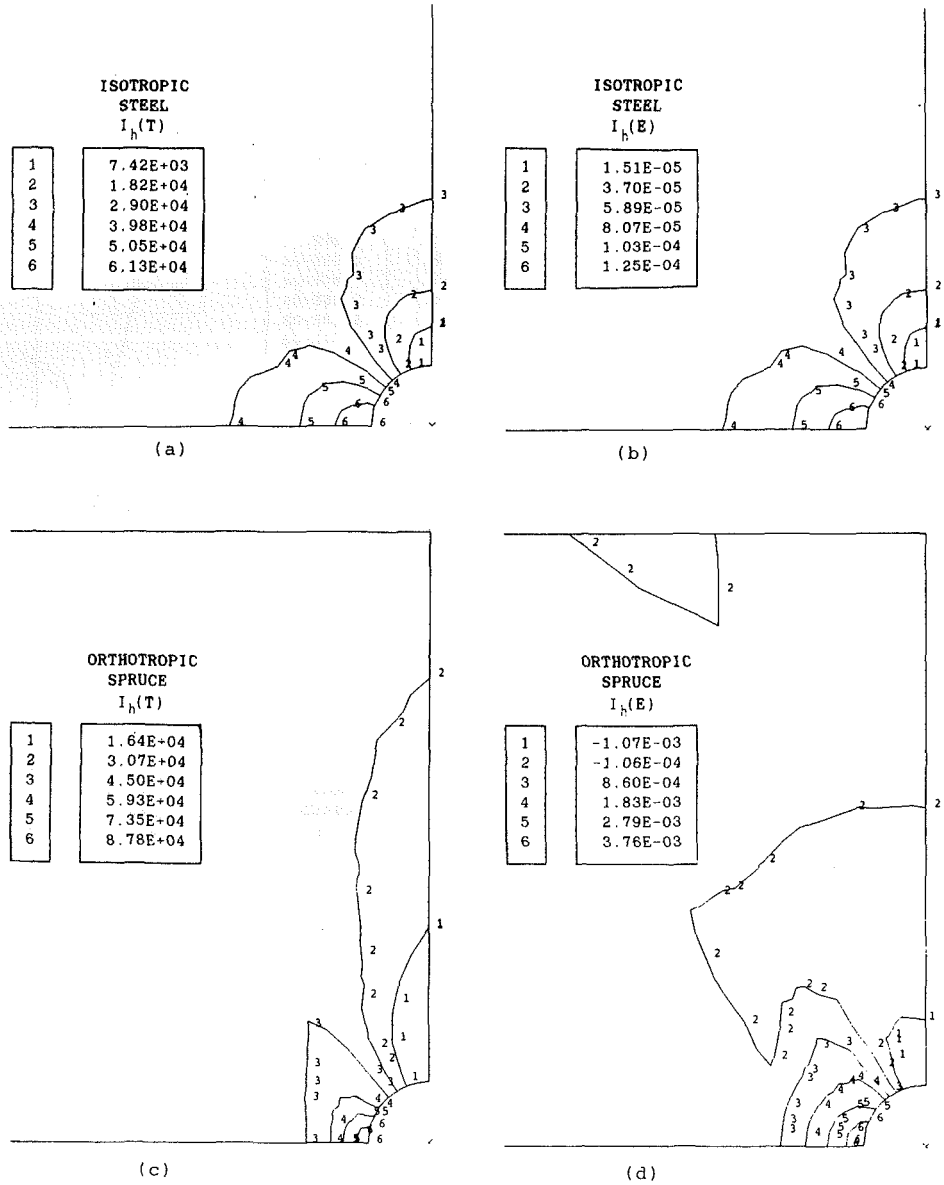


Fig. 5 The hydrostatic stress and strain for the steel and spruce examples: (*a*) hydrostatic stress, steel, (*b*) hydrostatic strain, steel, (*c*) hydrostatic stress, spruce, (*d*) hydrostatic strain, spruce

0.3527, 0.85, and 16.31, respectively. For the fourth, fifth, and sixth invariants, the proportionality constant is equal to the square of the corresponding eigenvalue of $\hat{\mathbf{c}}$; these proportionality constants are the squares of 1.22, 0.078, and 1.52, respectively.

The von Mises stress and strain of spruce are not proportional, however. This can be readily seen by comparing the contours of Fig. 4(*c*) to those of Fig. 4(*d*). On the other hand, the von Mises stress and strain of the isotropic steel are proportional. This can be readily seen by comparing the contours of Fig. 4(*a*) to those of Fig. 4(*b*).

Comparing the hydrostatic part of the stress tensor and the strain tensor of spruce, shown in Figs. 5(c) and 5(*d*), one can again easily see the lack of proportionality. The hydrostatic part of the stress and strain tensors of steel are proportional as seen from Figs. 5(*a*) and 5(*b*).

The objective of this paper is met if the reader has gained an appreciation of the importance of representing the stress and strain fields in linear anisotropic materials using contour plots of the proportional invariants rather than any other scalar valued function of the stress or strain. We have tried to demonstrate the significance of the proportional invariants by showing that, in the case of isotropic materials, this approach coincides with the traditional method of plotting the contours of the hydrostatic and von Mises stresses. The lack of proportionality of the hydrostatic and the von Mises stresses and strains for anisotropic materials has been demonstrated here. The results given here also suggest that the proportional invariants might be used in the development of the material strength or failure criteria.

Acknowledgment

This investigation was supported by NSF Grant No. BSC-8822401. This research was also supported (in part) by grant number 669301 from the PSC-CUNY Research Award Program of the City University of New York. We thank Prof. M. M. Mehrabadi for comments on an earlier draft of this manuscript. We also thank Mr. Yves Arramon for his assistance with the computations and the figures.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

Cowin, S. C., 1979, "On the Strength Anisotropy of Bone and Wood," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 832–838.

Cowin, S. C., 1990, "Deviatoric and Hydrostatic Mode Interaction in Hard and Soft Tissue," *Journal of Biomechanics*, Vol. 23, pp. 11–14.

Green, A. E., and Atkins, J. E., 1960, *Large Elastic Deformations and Non-Linear Continuum Mechanics*, Clarendon Press, Oxford, U.K.

Hill, R., 1950, *Plasticity*, Oxford University Press, Oxford, U.K.

Malmeisters, A. K., 1966, "Geometry of Theories of Strength," *Polymer Mechanics*, Vol. 2, p. 324.

Mehrabadi, M. M., and Cowin, S. C., 1990, "Eigentensors of Linear Anisotropic Elastic Materials," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. 43, pp. 15–41.

Paul, B., 1968, "Macroscopic Criteria for Plastic Flow and Brittle Fracture," *Fracture*, Academic Press, New York.

Rowlands, R. E., 1985, "Strength (Failure) Theories and Their Experimental Correlations," *Handbook of Composites, Vol. 3, Failure Mechanics of Composites*, Elsevier, Amsterdam, The Netherlands.

Tsai, S. W., and Wu, E. M., 1971, "A General Theory of Strength for Anisotropic Materials," *Journal of Composite Materials*, Vol. 5, p. 58.

von Mises, R., 1928, "Mechanik der plastichen Formandering von Kristallen," *Zeitschrift fur Angewandte Mathematik und Mechanik*, Vol. 8, pp. 161–185.

Wu, E. M., 1974, "Phenomenological Anisotropic Failure Criterion," *Composite Materials, Vol. 2, Mechanics of Composite Materials*, Academic Press, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stress-Focusing Effect in a Uniformly Heated Solid Sphere

Toshiaki Hata
Professor of Faculty of Education,
Shizuoka University,
Shizuoka City, 422, Japan

The ray theory is applied to the stress-focusing effects in a uniformly heated solid sphere. The stress-focusing effect is the phenomenon that, under an instantaneous heating, stress waves reflected from the free surface of the sphere result in very high stresses at the center. Using the ray theory, the Laplace transformed solution of stress waves in the sphere is sorted out into rays according to the ray path of multiply-reflected waves. Inverse transform of each ray gives rise to the exact solution of the transient response up to the arrival time of the next ray. The numerical results reveal that stresses peak out periodically at a constant period and, unlike the case of cylinder, the radial stress at the center of the sphere is bounded.

Introduction

The analysis of a long cylindrical rod subjected to a sudden rise in temperature uniformly over its cross-section has been studied by Ho (1976). This type of temperature distribution may be caused by the absorption of electromagnetic radiant energy or gamma ray of nuclear energy. One important conclusion by Ho (1976) is that, due to the instantaneous heating, stress waves reflected from the cylindrical surface of the rod may accumulate at the center and give rise to very high stresses, even though the initial thermal stress is relatively small. The same phenomenon may also be expected in a sphere, but has not been fully studied to the author's knowledge.

However, recently Hata (1991) solved the problem of thermal shock in a hollow sphere caused by a rapid uniform heating. In the study, he pointed out that if the ratio of the outer radius to the inner radius of the sphere increases, the peak tangential stress at the internal surface becomes higher and higher. As a limit case, he concluded that for a solid sphere, it might be possible to observe the stress-focusing effect. But we should note that the phenomenon in the aforementioned limit case of the hollow sphere problem is not the same as that in the solid sphere problem, because the boundary condition at the center of the sphere between two cases is different.

The results presented by Ho (1976) are obtained approximately by using the Laplace transform in time domain. The main difficulty in getting an exact solution by Laplace transform is the inversion of Laplace transform. Recently, by applying the ray theory, Hata (1991) shows that the complicated solutions of the thermoelastic problem in a spherical shell reduce to a simple form, which is suitable for the inverse Laplace transform. The ray theory derives its name from the geometric theory of light by following the propagation of a light pulse along a ray path, but differs from that in representing each path of the wave with an exact integral, known as the ray integral. These ray integrals can be either integrated in closed form or evaluated numerically with controlled accuracy.

In the paper, applying the ray theory to the problem of stress-focusing effect in a uniformly heated solid sphere, we find that the peaks of stresses due to the rapid heating become higher when the wave propagates toward the center of a sphere. But at the center of the sphere we do not observe the unbound stress peaks reported by Ho (1976) for the case of cylindrical rod.

Analysis

An isotropic, homogeneous, elastic solid sphere of radius b is subjected to a sudden uniform temperature rise over its cross-section. The radial displacement u is then the only nonzero displacement in the spherical coordinates (r, φ, θ).

The equation of motion for thermoelasticity is

$$\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} - 2\frac{u}{r^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} + \alpha\frac{(1+\nu)}{(1-\nu)}\frac{\partial T}{\partial r} \tag{1}$$

where α is the coefficient of thermal expansion and ν is Poisson's ratio and c is the dilatational wave speed, which is denoted as follows:

$$c = \sqrt{(\lambda + 2\mu)/\rho_0} \tag{2}$$

where λ and μ are the Lame's constants and ρ_0 is the density.

The temperature distribution in equation (1) is assumed to have the following form

$$\frac{T}{T_0} = \frac{1}{t_0}\{tH(t) - (t-t_0)H(t-t_0)\}. \tag{3}$$

The solid material is also assumed to be in the unstressed state

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 6, 1988; final revision Jan. 8, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and at a uniform temperature $T=0$ prior to time $t=0$. Hence, the initial conditions of the problem are taken as

$$u(r,t)=u(r,t)_{,t}=0 \qquad \text{at } t=0. \tag{4}$$

The boundary condition is that the radial displacement at $r=0$ vanishes and the radial stress σ_r satisfies the condition

$$\sigma_r(r,t)=0 \qquad \text{at } r=b. \tag{5}$$

The solution of equation (1), which satisfies the initial and boundary conditions, is obtained by using the method proposed by Hata (1989).

Applying the Laplace transform to equation (1) and invoking the fact that the temperature distribution of equation (3) is independent of r, the general solution of the transformed radial displacement is expressed in terms of spherical Hankel functions as follows:

$$\bar{u}(r,p)=C_1\{h_0^{(1)}(ipr/c)_{,r}+h_0^{(2)}(ipr/c)_{,r}\} \tag{6}$$

where C_1 is the unknown constant to be determined from the boundary condition, $h_0^{(1),(2)}(z)=\exp(\pm iz)/(\pm iz)$ are the zeroth-order spherical Hankel functions of the first and second kinds, respectively, p is the Laplace transform parameter, and overbars denote the Laplace transformed field.

Here we show that equation (6) vanishes at $r=0$. That is

$$\lim_{r\to 0}\{h_0^{(1)}(ipr/c)_{,r}+h_0^{(2)}(ipr/c)_{,r}\}$$

$$=-\lim_{r\to 0}(2c/p)\{\sinh(pr/c)-(pr/c)\cosh(pr/c)\}/r^2$$

$$=\lim_{r\to 0}(p/c)\sinh(pr/c)=0. \tag{7}$$

It is seen from equation (7) that equation (6) satisfies the condition of zero radial displacement at the center of the sphere. Stresses resulting from the temperature field of equation (3) are

$$\sigma_r(r,p)=-\frac{\epsilon T_0\rho_0 c^2}{t_0}\frac{1}{p^2}(1-e^{-pt_0})$$

$$+C_1\left\{-\frac{4\mu}{r}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}r\right)+\rho_0 p^2 h_0^{(1)}\left(i\frac{p}{c}r\right)\right\}$$

$$+C_1\left\{-\frac{4\mu}{r}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}r\right)+\rho_0 p^2 h_0^{(2)}\left(i\frac{p}{c}r\right)\right\}$$

$$\sigma_\theta(r,p)=-\frac{\epsilon T_0\rho_0 c^2}{t_0}\frac{1}{p^2}(1-e^{-pt_0})$$

$$+C_1\left\{\frac{2\mu}{r}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}r\right)+\frac{\lambda p^2}{c^2}h_0^{(1)}\left(i\frac{p}{c}r\right)\right\}$$

$$+C_1\left\{\frac{2\mu}{r}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}r\right)+\frac{\lambda p^2}{c^2}h_0^{(2)}\left(i\frac{p}{c}r\right)\right\} \tag{8}$$

where

$$\epsilon=\{(1+\nu)/(1-\nu)\}\alpha \tag{9}$$

and $h_0^{(1),(2)\prime}(z)=dh_0^{(1),(2)}(z)/dz$.

The unknown constant C_1 can be determined by the requirement of zero radial stress, equation (5), at the surface of the sphere. The constant C_1 is determined as

$$C_1=-\bar{p}_o/(C_{21}+C_{22}) \tag{10}$$

where

$$\bar{p}_o=-\frac{\epsilon T_o\rho_o c^2}{t_o}\frac{(1-e^{-pt_o})}{p^2}$$

$$C_{21}=-\frac{4\mu}{b}\left(i\frac{p}{c}\right)h_o^{(1)\prime}\left(i\frac{p}{c}b\right)+\rho_o p^2 h_o^{(1)}\left(i\frac{p}{c}b\right)$$

$$C_{22}=-\frac{4\mu}{b}\left(i\frac{p}{c}\right)h_o^{(2)\prime}\left(i\frac{p}{c}b\right)+\rho_o p^2 h_o^{(2)}\left(i\frac{p}{c}b\right). \tag{11}$$

The Ray Integrals

The solution given by equations (6) and (10) can be cast into a different form by introducing a reflection coefficient R, which is defined

$$R=-\frac{\bar{C}_{21}}{C_{22}}=\left(\frac{p^2b^2/c^2+4\kappa pb/c+4\kappa}{p^2b^2/c^2-4\kappa pb/c+4\kappa}\right)e^{-2\frac{pb}{c}} \tag{12}$$

where $\kappa=\mu/(\rho_0 c^2)$. The function R is the reflection coefficient for an outgoing spherical wave being reflected at the traction-free surface $r=b$. Note that, unlike the case of a hollow sphere (Hata, 1990), the solid sphere has only one reflection coefficient. Hence, we have from equation (6) that

$$\bar{u}(r,p)=-\frac{\bar{p}_o}{C_{22}(1-R)}\left\{h_o^{(1)}(ipr/c)_{,r}+h_o^{(2)}(ipr/c)_{,r}\right\}$$

$$=-\frac{\bar{p}_o}{C_{22}}(1+R+R^2+\ldots)\{h_o^{(1)}(ipr/c)_{,r}+h_o^{(2)}(ipr/c)_{,r}\}$$

$$\times(|R|<1). \tag{13}$$

In general, the condition of $|R|<1$ can be met by choosing an appropriate integration path for the inverse Laplace transform. Equation (13) can be written as

$$\bar{u}(r,p)=\sum_{j=0}^{\infty}\bar{\psi}_j(r,p) \tag{14}$$

where

$$\bar{\psi}_o(r,p)=-\frac{\bar{p}_o}{C_{22}}h_o^{(2)}(ipr/c)_{,r}$$

$$\bar{\psi}_1(r,p)=-\frac{\bar{p}_o}{C_{22}}h_o^{(1)}(ipr/c)_{,r}$$

$$\bar{\psi}_j(r,p)=-R(p)\bar{\psi}_{j-2}(r,p) \qquad (j=2,3,\ldots) \tag{15}$$

Each term in the series in equation (14) represents a spherical wave traveling in the medium. The function $\bar{\psi}_o$ is an incoming wave, which is generated by an imaginary uniform pressure $-\bar{p}_o$ of equation (11) applied at the cross-section of the sphere; $\bar{\psi}_1$ is an outgoing wave which is generated by the center of the sphere $r=0$. Similarly, $\bar{\psi}_{2j}$ represents the incoming waves, while $\bar{\psi}_{2j+1}$ represents the outgoing waves. Thus, $\bar{\psi}_j$ represents the multiply-reflected waves between the center and the surface of the sphere. The corresponding stresses of equation (8) can be expressed similarly in a series of rays.

The inverse transform of each term then represents the transient waves which are continuously reflected between the center and the surface. The complete solution is

$$u(r,t)=\sum_{j=0}^{\infty}\frac{1}{2\pi i}\int_{\epsilon_0-i\infty}^{\epsilon_0+i\infty}\bar{\psi}_j(r,p)e^{pt}dp. \tag{16}$$

We call each term of the series a ray integral of the transient wave solution (Pao and Ceranoglu, 1978).

Transient Solution

It is seen from equation (15) that, once the inversions of $\bar{\psi}_0$ and $\bar{\psi}_1$ are carried out, the higher-order terms in the series can be obtained by a convolution. Denoting the inverse transform of $\bar{\psi}_j(r,p)$ by $\psi_j(r,t)$, we find from equation (15) that

$$\psi_j(r,t)=R(t)*\psi_{j-2}(r,t)=\int_o^t R(\tau)\psi_{j-2}(r,t-\tau)d\tau. \tag{17}$$

The inverse Laplace transform of $R(p)$ in equation (12) is given by

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$R(t)=\frac{1}{2\pi i}\int_{\epsilon_o-i\infty}^{\epsilon_o+i\infty}\frac{(pb/c+\gamma)(pb/c+\gamma^*)}{(pb/c-\gamma)(pb/c-\gamma^*)}e^{pt-2\frac{pb}{c}}dp \quad (18)$$

where

$$\gamma=2\kappa-2\sqrt{\kappa^2-\kappa},\quad \gamma^*=2\kappa+2\sqrt{\kappa^2-\kappa} \quad (19)$$

and $\kappa=\mu/(\lambda+2\mu)$.

Using the residue theorem, we find

$$R(t)=\delta\left(t-\frac{2b}{c}\right)+\sum_{i=1}^{2}\frac{4\kappa p_i}{(p_ib/c)-2\kappa}e^{p_i(t-2b/c)}H\left(t-\frac{2b}{c}\right) \quad (20)$$

where $\delta(t)$ is a delta function and the two simple poles $p_i(i=1,2)$ are

$$p_1=c\gamma/b,\quad p_2=c\gamma^*/b \quad (21)$$

Inverse Laplace transforms of $\overline{\psi_0}(r,p)$ and $\overline{\psi_1}(r,p)$ can be obtained in closed form by a contour integration similar to that for $R(t)$. Hence, we obtain from equation (15)

$$\psi_o(r,t)=-\left(\frac{\epsilon T_o b^3}{t_o r^2}\right)\left[\frac{t_o}{\gamma\gamma^*}H\left(t+\frac{r}{c}-\frac{b}{c}\right)+\sum_{i=1}^{2}\left(1-\frac{p_ir}{c}\right)\left\{e^{p_i\left(t+\frac{r}{c}-\frac{b}{c}\right)}H\left(t+\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0+\frac{r}{c}-\frac{b}{c}\right)}H\left(t-t_0+\frac{r}{c}-\frac{b}{c}\right)\right\}/p_ig_j\right] \quad (22)$$

$$\psi_1(r,t)=\left(\frac{\epsilon T_o b^3}{t_o r^2}\right)\left[\frac{t_0}{\gamma\gamma^*}H\left(t-\frac{r}{c}-\frac{b}{c}\right)+\sum_{i=1}^{2}\left(1+\frac{p_ir}{c}\right)\left\{e^{p_i\left(t-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-t_0-\frac{r}{c}-\frac{b}{c}\right)\right\}/p_ig_j\right] \quad (23)$$

where

$$g_i=4(p_ib/c)^2-12\kappa(p_ib/c)+8\kappa. \quad (24)$$

The high-order terms in the series are obtained through the convolution integral of equation (17). Substitution of equations (20) into (17) yields, for $j\geqq 1$,

$$\psi_{2j}(r,t)=\int_o^t\left\{\delta\left(\tau-\frac{2b}{c}\right)+R_c(\tau)\right\}\psi_{2(j-1)}(r,t-\tau)d\tau$$
$$=\psi_{2(j-1)}\left(r,t-\frac{2b}{c}\right)+H\left(t+\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\times\int_{2b/c}^{t+r/c-b/c-2(j-1)(b/c)}R_c(\tau)\psi_{2(j-1)}(r,t-\tau)d\tau \quad (25)$$

$$\psi_{2j+1}(r,t)=\int_o^t\left\{\delta\left(\tau-\frac{2b}{c}\right)+R_c(\tau)\right\}\psi_{2j+1}(r,t-\tau)d\tau$$
$$=\psi_{2j-1}\left(r,t-\frac{2b}{c}\right)+H\left(t-\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\times\int_{2b/c}^{t-r/c-b/c-2(j-1)(b/c)}R_c(\tau)\psi_{2j-1}(r,t-\tau)d\tau \quad (26)$$

where

$$R_c(t)=\sum_{i=1}^{2}\frac{4\kappa p_i}{p_ib/c-2\kappa}e^{p_i(t-2b/c)}H\left(t-\frac{2b}{c}\right). \quad (27)$$

Using equations (25) and (26) and the conditions of

$$\psi_{2(j+1)}(r,-r/c+b/c+2(j-1)b/c)=\psi_{2j-1}(r,r/c+b/c+2(j-1)b/c)=0,$$

we obtain

$$\psi_{2j}(r,t),_r=\psi_{2(j-1)}(r,t-2b/c),_r+H\left(t+\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\times\int_{2b/c}^{t+r/c-b/c-2(j-1)(b/c)}R_c(\tau)\psi_{2j-2}(r,t-\tau),_r d\tau \quad (28)$$

$$\psi_{2j+1}(r,t),_r=\psi_{2j-1}(r,t-2b/c),_r+H\left(t-\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\times\int_{2b/c}^{t-r/c-b/c-2(j-1)(b/c)}R_c(\tau)\psi_{2j-1}(r,t-\tau),_r d\tau. \quad (29)$$

The displacement u is derived from equation (14) as

$$u(r,t)=\sum_{j=0}^{\infty}\psi_j(r,t). \quad (30)$$

Using the stress-strain relations, we obtain the corresponding stresses as follows:

$$\sigma_r^E(r,t)=\rho_oc^2[u,_r+\{2\nu/(1-2\nu)\}(u/r)]$$
$$\sigma_\theta^E(r,t)=\rho_oc^2[\{\nu/(1-\nu)\}u,_r+\{1/(1-\nu)\}(u/r)]. \quad (31)$$

Finally, the thermal stresses are

$$\sigma_r=\sigma_r^E+\sigma^T$$
$$\sigma_\theta=\sigma_\theta^E+\sigma^T \quad (32)$$

where σ^T is given by

$$\sigma^T=-\epsilon\rho_oT_oc^2\{tH(t)-(t-t_o)H(t-t_o)\}/t_o. \quad (33)$$

Next, we try to determine the behavior at the center of the sphere. First, we show that the displacement, as shown in equation (30), at the center is zero. This has already been shown for the Laplace transformed displacement of equation (7). In equations (22) and (23) it seems that each of ψ_0 and ψ_1 becomes infinity at the limit of $r\to 0$. However, the sum of two functions $\psi_0+\psi_1$, which is the sum of equations (22) and (23),

$$\psi_o+\psi_1=\left(\frac{\epsilon T_o b^3}{t_o r^2}\right)\left[\frac{t_o}{\gamma\gamma^*}\left\{-H\left(t+\frac{r}{c}-\frac{b}{c}\right)+H\left(t-\frac{r}{c}-\frac{b}{c}\right)\right\}-\sum_{i=1}^{2}\left(1-\frac{p_ir}{c}\right)\left\{e^{p_i\left(t+\frac{r}{c}-\frac{b}{c}\right)}H\left(t+\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0+\frac{r}{c}-\frac{b}{c}\right)}H\left(t-t_0+\frac{r}{c}-\frac{b}{c}\right)\right\}+\sum_{i=1}^{2}\times\left(1+\frac{p_ir}{c}\right)\left\{e^{p_i\left(t-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-t_0-\frac{r}{c}-\frac{b}{c}\right)\right\} \quad (34)$$

vanishes at the center of the sphere, because

$$\lim_{r\to 0}(\psi_o+\psi_1)=\lim_{r\to 0}\sum_{i=1}^{2}-\left(\frac{2\epsilon T_o b^3}{t_o}\right)\times\frac{\{\sinh(p_ir/c)-(p_ir/c)\cosh(p_ir/c)\}}{r^2}\left\{e^{p_i(t-b/c)}H\left(t-\frac{b}{c}\right)-e^{p_i(t-t_0-b/c)}H\left(t-t_0-\frac{b}{c}\right)\right\}/p_ig_i=0. \quad (35)$$

Similarly it can be shown that $\psi_{2j}+\psi_{2j+1}$ vanishes at the center

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of the sphere. Therefore, the displacement at the center is zero, which is one of the boundary conditions.

Next, we try to show that the stresses at the center of the sphere are bounded. We have from equation (34) that

$$\psi_{0,r}+\psi_{1,r}=\left(\frac{\epsilon T_o b^3}{t_o r^3}\right)\left[2\frac{t_o}{\gamma\gamma^*}\left\{H\left(t+\frac{r}{c}-\frac{b}{c}\right)-H\left(t-\frac{r}{c}-\frac{b}{c}\right)\right\}+\sum_{i=1}^{2}\left\{2-2\frac{p_i r}{c}+\left(\frac{p_i r}{c}\right)^2\right\}\times\left\{e^{p_i\left(t+\frac{r}{c}-\frac{b}{c}\right)}H\left(t+\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0+\frac{r}{c}-\frac{b}{c}\right)}\times H\left(t-t_o+\frac{r}{c}-\frac{b}{c}\right)\right\}-\sum_{i=1}^{2}\left\{2+2\frac{p_i r}{c}+\left(\frac{p_i r}{c}\right)^2\right\}\times\left\{e^{p_i\left(t-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-\frac{r}{c}-\frac{b}{c}\right)-e^{p_i\left(t-t_0-\frac{r}{c}-\frac{b}{c}\right)}H\left(t-t_o-\frac{r}{c}-\frac{b}{c}\right)\right\}. \quad (36)$$

By taking the limit $r\to 0$, we get

$$\lim_{r\to 0}(\psi_{0,r}+\psi_{1,r})=\lim_{r\to 0}\left(\frac{2\epsilon T_0 b^3}{t_0}\right)\sum_{i=1}^{2}\times\frac{2\left(\frac{p_i}{c}\right)^3 r\cosh\left(\frac{p_i}{c}r\right)+\left(\frac{p_i}{c}\right)^4 r^2\sinh\left(\frac{p_i}{c}r\right)}{6r}\times\left\{e^{p_i\left(t-\frac{b}{c}\right)}H\left(t-\frac{b}{c}\right)-e^{p_i\left(t-t_0-\frac{b}{c}\right)}H\left(t-t_0-\frac{b}{c}\right)\right\}=\frac{2\epsilon T_0 b^3}{3t_0}\sum_{i=1}^{2}\left(\frac{p_i}{c}\right)^3\left\{e^{p_i\left(t-\frac{b}{c}\right)}H\left(t-\frac{b}{c}\right)-e^{p_i\left(t-t_0-\frac{b}{c}\right)}H\left(t-t_0-\frac{b}{c}\right)\right\}. \quad (37)$$

The sum of equations (28) and (29) yields

$$\psi_{2j,r}+\psi_{2j+1,r}=\psi_{2(j-1),r}+\psi_{2j-1,r}+H\left(t+\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\times\int_{2b/c}^{t+r/c-b/c-2(j-1)(b/c)}R_c(\tau)\psi_{2j-2}(r,t-\tau),_r d\tau+H\left(t-\frac{r}{c}-\frac{b}{c}-2j\frac{b}{c}\right)\int_{2b/c}^{t-r/c-b/c-2(j-1)(b/c)}\times R_c(\tau)\psi_{2j-1}(r,t-\tau),_r d\tau. \quad (38)$$

Taking $r\to 0$ in equation (38), we have

$$\lim_{r\to 0}(\psi_{2j,r}+\psi_{2j+1,r})=\{\psi_{2j-1,r}+\psi_{2(j-1),r}\}/_{r=0}+H\left(t-\frac{b}{c}-2j\frac{b}{c}\right)\int_{2b/c}^{t-b/c-2(j-1)(b/c)}\times R_c(\tau)\{\psi_{2j-1}(r,t-\tau),_r+\psi_{2(j-1)}(r,t-\tau),_r\}/_{r=0}d\tau. \quad (39)$$

If $j=1$, substituting equation (37) and $R_c(r)$ of equation (27) into equation (39), we obtain $\psi_{2,r}+\psi_{3,r}$ at the center. Since $\psi_{0,r}+\psi_{1,r}$ at the center in equation (37) is bounded and $R_c(t)$ is not a function of r, it is easily seen that equation (39) is bounded. The function $\psi_{4,r}+\psi_{5,r}$, which is obtained by the same procedure as before, is also bounded. Thus, the function $\psi_{2j,r}+\psi_{2j+1,r}$, which is determined successively by the foregoing procedure, must be bounded. Therefore, from equations (31), (35), (37), and (39) it is concluded that the stresses at the center of the sphere are bounded unlike the cylindrical rod case, in which the stresses at the center are unbounded (Ho, 1976).

Results and Discussions

To show the stress-focusing effect in a uniform heated solid sphere, we performed numerical calculations for a sphere with Poisson's ratio $\nu=0.3$. In an instantaneous heating case of $t_0=0$, the temperature distribution in equation (1) becomes $T=T_0H(t)$. The numerical integrations were carried out using Simpson's formula. The numerical results are shown in Figs. 1 to 4(b). In these figures the nondimensional variables t^* and σ_i^* are defined as $t^*=tc/b$ and $\sigma_i^*=\sigma_i/(\epsilon T_0\rho_0 c^2)$ $(i=r,\theta)$.

The values of individual first nine rays of σ_r^{E*} at the surface of the sphere are shown in Fig. 1. Since the stress σ_r^{E*} at the surface must be one in order to delete the initial stress $\sigma_r^{T*}=-1$, the stress caused by the ray cancels each other and the sum of rays in Fig. 1 tends to be one. Then the sum of σ_r^{E*} and σ_r^{T*} satisfies the stress-free boundary condition on the surface of the sphere. Here it is noted that in equations (22) and (23) the step functions define the arrival times of each of the rays, which arrive in successive order. The first ray, ψ_0, arrives at $t^*=0$. The second ray, ψ_1, arrives at $t^*=2$. Similarly, rays ψ_2 and ψ_3 arrive, respectively, at $t^*=2$ and 4. Since high-order rays all arrive at a later time, the solution obtained by summing the first j rays is exact up to the time of arrival of the $(j-1)$th ray.

The behavior of radial stress as a function of time is illustrated in Fig. 2 at the locations of $r/b=0$, 0.01, 0.1, 0.5, and 1.

In Fig. 2 the stress σ_r^* is initially compressive because of the instantaneous heating. After this period, the waves reflect from the surface and accumulate to the center. The maximum value of σ_r^* at $r/b=0.1$ is about four times higher than that at $r/b=0.5$. As the waves approach to the center, the maximum values of σ_r^* increase higher and higher as shown in Fig. 2. The peak of stress σ_r^* at $r/b=0.01$, which is about five times higher than that at $r/b=0.01$, peaks out periodically at an interval of $t^*=2$. But it is noticed that the stress σ_r^* at the center has no peak as mentioned in equations (36) to (39).

Figure 3 shows the variations of tangential stress σ_θ^* at different r/b values as a function of time. The behavior of tangential stress σ_θ^* is similar to that of the radial stress. In the figures we also observe that the stress occurred near the center

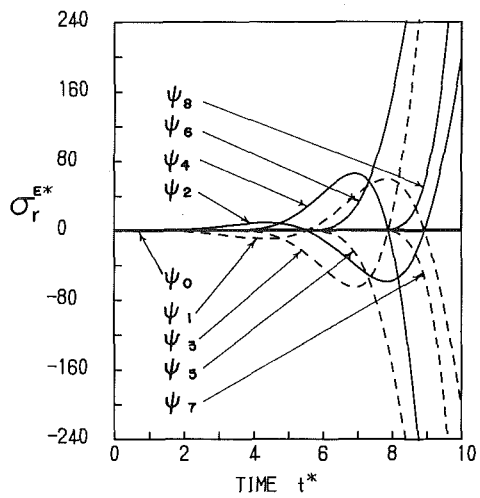


Fig. 1 Values of σ_r^{E*} as a function of time for individual rays on the surface ($r/b=1$)

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

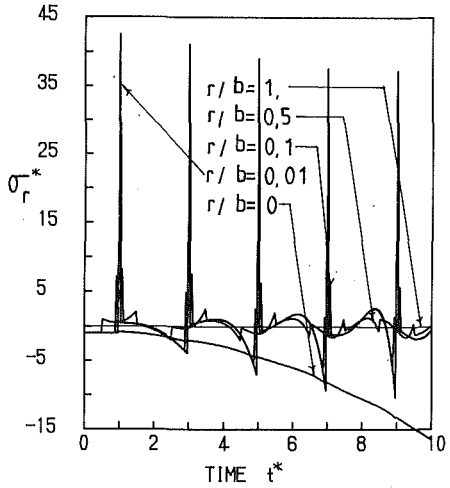


Fig. 2 Stress-focusing effect of the radial stress σ_r^*

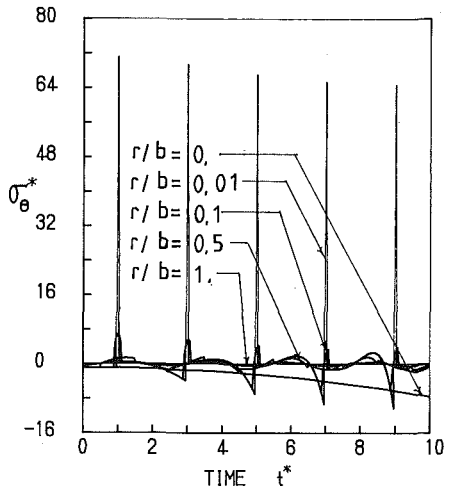


Fig. 3 Stress-focusing effect of the tangential stress σ_θ^*

peaks out periodically. It is important that the high-peak stresses in the sphere are taken into account in the design of nuclear reactor.

Here, we should explain the reason why the stress response changes from the curve with periodical peaks near $r=0$ to the smooth curve at $r=0$ as shown in Figs. 2 and 3. The values of individual first ten rays of σ_r^{E*} at $r/b=0.01$ are shown in Fig. 4(*a*) and the values of the sum of two rays, $\psi_{2j}+\psi_{2j+1}$ ($j=0$ to 4), of σ_r^{E*} at $r/b=0.01$ are shown in Fig. 4(*b*). In Fig. 4(*a*) the first ray, ψ_0, arrives at $t^*=0.99$ at $r/b=0.01$, obtained from equation (22), and is diverging in time t^*. The second ray, ψ_1, arrives at $t^*=1.01$ at $r/b=0.01$, obtained from equation (23), and is also diverging in time t^*. The other rays are individually diverging in time. However, in Fig. 4(*b*) we can observe that the sum of first and second rays, $\psi_0+\psi_1$, of the stress σ_r^{E*} shows the peak in the time interval between $t^*=0.99$ and 1.01 and, after the arrival of the second ray, a pair of rays are combined and the sum is convergent. The same phenomena can be observed in other curves in Fig. 4(*b*). Therefore, the radial stress σ_r^* at $r/b=0.01$ peaks out periodically as shown in Fig. 2. In addition, since the arrival time of the ray, ψ_{2j}, is the same as that of the ray, ψ_{2j+1}, at the center of the sphere, the curve at $r/b=0$ has no peak as shown in Fig. 2.

Next, we discuss the behavior of stress at the center of the sphere in precise time. In the instantaneous heating case of $t_0=0$ we derive from equation (37)

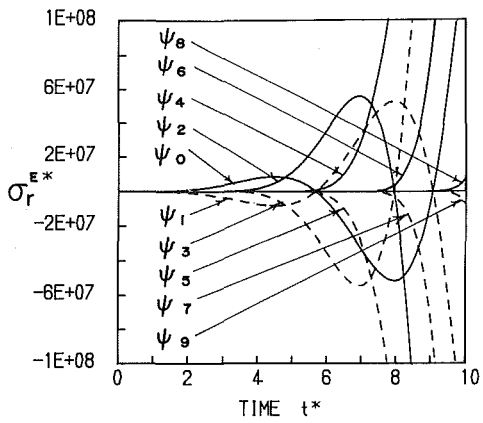


Fig. 4(*a*) Values of σ_r^{E*} as a function of time for individual rays at $r/b=0.01$

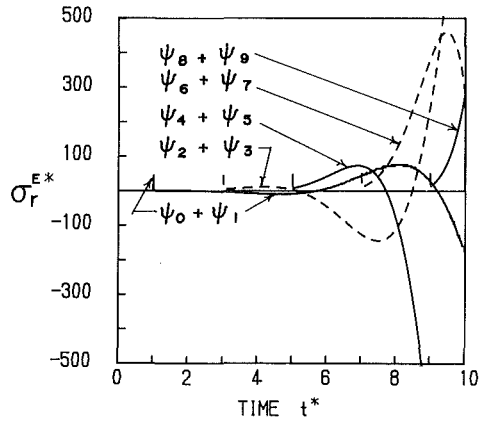


Fig. 4(*b*) Values of σ_r^{E*} as a function of time for sums of two rays at $r/b=0.01$

$$\lim_{r\to 0}(\psi_{0,r}+\psi_{1,r})=\frac{2\epsilon T_o b^3}{3}\sum_{i=1}^{2}\left(\frac{p_i}{c}\right)^3 p_i e^{p_i\left(t-\frac{b}{c}\right)}H\left(t-\frac{b}{c}\right). \tag{40}$$

Substitution of the two poles in equation (21) into equation (40) yields

$$\lim_{r\to 0}(\psi_{0,r}+\psi_{1,r})=\frac{64}{3}\frac{\epsilon T_o c}{b}e^{2\kappa\frac{c}{b}\left(t-\frac{b}{c}\right)}\kappa^2\Bigg[(8\kappa^2-8\kappa+1)\cosh\left\{2\sqrt{\kappa^2-\kappa}\left(\frac{ct}{b}-1\right)\right\}+4(2\kappa-1)\sqrt{\kappa^2-\kappa}\sinh\left\{2\sqrt{\kappa^2-\kappa}\left(\frac{ct}{b}-1\right)\right\}\Bigg]H\left(t-\frac{b}{c}\right). \tag{41}$$

Since Poisson's ratio ν takes the value of $0\leqq\nu\leqq 0.5$, κ varies from 0 to 0.5. If we substitute $\nu=0.5$, where $\kappa=0$, into equation (41), the sum $\psi_{0,r}+\psi_{1,r}$ in equation (41) equals to zero. Hence, the stresses at the center of the sphere are zero. If we substitute $\nu=0$, where $\kappa=0.5$, into equation (41), the stresses at the center of the sphere are given by using the equation

$$\lim_{r\to 0}(\psi_{0,r}+\psi_{1,r})=-\frac{16\epsilon T_o c}{3b}e^{\frac{c}{b}\left(t-\frac{b}{c}\right)}\cos\left(\frac{ct}{b}-1\right)H\left(t-\frac{b}{c}\right). \tag{42}$$

The value of equation (42) tends to be infinity as t^* approaches

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

infinity. Taking into account of equation (39), the stresses at the center of the sphere also tend to be infinity as t^* approaches infinity. In this paper, since we taken Poisson's ratio $\nu = 0.3$, where $\kappa = 0.29$, and substitute it into equations (41) and (39), the stresses at the center of the sphere may be divergent as time becomes larger. Therefore, the curves at $r/b = 0$ in Figs. 2 and 3 seem to be divergent as t^* goes to infinity.

It should be noted that from the foregoing discussion we derived the stresses at the center for time $t = b/c$ are bounded, but Mann-Nachbar and Nachbar (1970) discussed the singular behavior of both stresses and displacement at the center for time $t = b/c$ precisely and concluded that the Dirac delta function at time $t = b/c$ must be added to the right-hand side of equation (37).

Conclusion

The major accomplishment of this study has been in gaining a better understanding of the stress-focusing effect in a uniformly heated solid sphere. The stress-focusing effects occur when stress waves reflect from the surface of the sphere and accumulate radially inward to the center. The results presented are obtained by use of the Laplace transform on time. The transformed solution of the stress waves for the sphere is sorted out into rays according to the ray path of multiply reflected waves. Inverse transform of each ray gives rise to the exact solution of the transient response up to the arrival time of the next ray. Therefore, we could find the interesting phenomena that the stresses caused by the stress-focusing effects peak out periodically at a constant interval, and the stresses at the center of the sphere are bounded rather than unbounded as in the case of cylinder.

Acknowledgments

The author is especially grateful to Professor Emeritus A. Atsumi of Tohoku University for his invaluable suggestions. This work was partly supported by Grant-Aid of the National Science Foundation from the Ministry of Education in Japan.

References

Hata, T., 1991, "Thermal Shock in a Hollow Sphere caused by Rapid Uniform Heating," ASME Journal of Applied Mechanics, Vol. 58, pp. 64–69.

Ho, C. H., 1976, "Stress-Focusing Effect in a Uniformly Heated Cylindrical Rod," ASME Journal of Applied Mechanics, Vol. 43, pp. 404–468.

Mann-Nachbar, P., and Nachbar, W., 1970, "Thermal Shock Following Rapid Uniform Heating of Spheres and Long Cylindrical Rods," Aerospace Corp. Report TR-0066 (S5816-71)-1.

Pao, Y. H., and Ceranoglue, A. N., 1978, "Determination of Transient Response of a Thick-Walled Spherical Shell by Ray Theory," ASME Journal of Applied Mechanics, Vol. 45, pp. 114–122.

Takeuti, Y., 1971, *Analyses of Thermal Stresses*, Nisshin Syuppan (in Japanese).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Thermal Shock in a Hollow Sphere Caused by Rapid Uniform Heating

Toshiaki Hata
Professor of the Faculty of Education,
Shizuoka University,
Shizuoka City, 422, Japan

This paper is concerned with a method for calculating the dynamic stress distribution in a hollow sphere. Adopting the Goodier's concept, the dynamic thermoelastic problem in a sphere is decomposed into a particular form of dynamic stress problem corresponding to the thermoelastic displacement potential and the homogeneous form of dynamic stress problem corresponding to the stress functions. Applying the ray theory to the homogeneous form, we obtain the general solution for transient waves induced by sudden heating. When a hollow sphere is subjected suddenly to a uniform temperature rise throughout the sphere, stress waves occur at the internal and external surfaces the moment thermal impact is applied. During instantaneous heating, the interfering effects of these waves can cause a very high dynamic stress at the internal surface of a sphere. The numerical results show that, as the ratio of the internal radius to the external one increases, the tangential stress on the internal surface becomes higher.

Introduction

Consider a hollow spherical shell subjected to a sudden rise in temperature uniformly throughout the shell. In cases of sudden cooling or heating of a body, the inertia forces must be taken into account. Mathematical difficulties in treating the dynamic problem are considerable and, therefore, rigorous solutions have been obtained in a closed form for only a few simple one-dimensional problems. An investigation of dynamic "noncoupled" spherical symmetric problems was originated by Sternberg and Chakrovorty (1959), who considered the state of stress due to sudden heating of a spherical cavity.

As for the study of thick-walled spherical shells, Tsui and Kraus (1965) obtained the closed-form solution for thermal stress-wave propagation in hollow elastic spheres by using Laplace transformation. This solution, however, is too complex to provide a clear understanding of the wave propagation. Zaker (1966, 1968) presented a comprehensive analysis of the dynamic thermal shock in hollow spheres. His solution is based on the method of characteristics. The accuracy of the solution depends on the mesh size used for the numerical integration. The numerical results of Tsui, Kraus, and Zaker are exhibited only for the early-time wave propagation, such as the first or the second wave propagation.

Recently, various special ceramics are essential to the development of computers and many other electronic devices. When ceramic materials are subjected to a rapid change in temperature, such as thermal shock, substantial stresses develop because of temperature changes. Resistance to weakening or fracture under these conditions is called thermal endurance, thermal stress resistance, or thermal shock resistance. The effect of thermal stresses on different kinds of materials depends not only on material characteristics but also on stress level, stress distribution in the body, and stress duration. Hence, it is very important to determine the actual stress in a ceramic material under given conditions of heat transfer.

When a elastic hollow sphere is subjected suddenly to a uniform temperature rise throughout the sphere, stress waves occur at the internal and external surfaces the moment thermal impact is applied. The stress wave at the internal surface proceeds radially outward, while the wave at the external surface propagates radially inward. This type of temperature distribution may be developed by the absorption of infrared rays radiation or electromagnetic radiant energy from pulses that are typically of duration much less than 1μ sec.

In this paper we solve exactly the problem of thermal shock in a hollow sphere. In the analysis, adopting the Goodier's concept (Goodier, 1937), we find that the thermoelastic dynamic problem is decomposed into a particular form of dynamic stress problem solved by using the thermoelastic displacement potential and the homogeneous form of dynamic stress problem solved by using the stress functions. Applying the ray theory (Pao and Ceranoglue, 1978) to the homogeneous form of the problem, we obtain a general solution for transient wave in a thick-walled spherical shell generated by sudden heating. The numerical results show that the waves can cause very high dynamic stresses at the internal surface of a hollow sphere. These stresses increase a lot, as the ratio of the internal radius to the external radius increases.

Formulation of the Problem

The constitutive equations and field equations for a solid with thermoelastic deformation are presented in this section.

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 6, 1988; final revision, Oct. 31, 1988.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The solution of the problem requires the determination of the dynamic stress tensor, $\mathbf{T}$, the strain tensor, $\mathbf{S}$, and the displacement vector, $\mathbf{u}$, in the region Σ bounded by two boundaries Γ_1 and Γ_2. In the case of a sudden heating or cooling of the body we have to investigate the equations of motion as follows:

$$\nabla \cdot \nabla \mathbf{u} + \frac{1}{1-2\nu} \nabla \nabla \cdot \mathbf{u} = \frac{\rho_0}{\mu} \ddot{\mathbf{u}} + 2 \frac{1+\nu}{1-2\nu} \nabla \cdot \alpha T \mathbf{I} \quad (1)$$

where ν, μ, α, ρ_0 represent the Poisson's ratio, the shear modulus, the coefficient of thermal expansion, and the density, respectively; and T is temperature and $\mathbf{I}$ denotes the unit tensor.

Total displacement $\mathbf{u}$ consists of the thermoelastic displacement $\mathbf{u}^T$ and the elastic displacement $\mathbf{u}^E$. From equation (1) we have

$$\nabla \cdot \nabla \mathbf{u}^T + \frac{1}{1-2\nu} \nabla \nabla \cdot \mathbf{u}^T = \frac{\rho_0}{\mu} \ddot{\mathbf{u}}^T + 2 \frac{1+\nu}{1-2\nu} \nabla \cdot \alpha T \mathbf{I} \quad (2)$$

$$\nabla \cdot \nabla \mathbf{u}^E + \frac{1}{1-2\nu} \nabla \nabla \cdot \mathbf{u}^E = \frac{\rho_0}{\mu} \ddot{\mathbf{u}}^E. \quad (3)$$

The solution of equation (2) may be simply derived by using the thermoelastic displacement potential (Goodier, 1937). That is

$$\mathbf{u}^T = \nabla \phi^T. \quad (4)$$

Substituting equation (4) into (2), we know the system of equation (2) is satisfied if

$$\nabla \cdot \nabla \phi^T - \frac{1-2\nu}{2(1-\nu)} \frac{\rho_0}{\mu} \ddot{\phi}^T = \frac{1+\nu}{1-\nu} \alpha T. \quad (5)$$

The corresponding stresses may be written as

$$\mathbf{T}^T = 2\mu(\nabla \nabla \phi^T - \nabla \cdot \nabla \phi^T \mathbf{I}) + \rho_0 \ddot{\phi}^T \mathbf{I}. \quad (6)$$

For a bounded body the function ϕ^T satisfies only a part of the boundary conditions. Therefore, the incomplete solution of equation (5) must be completed by a solution $\mathbf{u}^E$ of the system of homogeneous equations of motion (3). The function $\mathbf{u}^E$ must so be chosen such that the final solution $\mathbf{u} = \mathbf{u}^T + \mathbf{u}^E$ satisfies all the boundary conditions of the problem.

For example, for a body with the zero surface tractions, the boundary conditions read

$$(\mathbf{T}^T + \mathbf{T}^E) \cdot \mathbf{n}_i = \mathbf{0} \quad \text{on } \Gamma_i \quad (i = 1, 2) \quad (7)$$

where $\mathbf{n}_i$ is the vector of surface normal.

Next, we express the components of the displacement vector $\mathbf{u}^E$ in terms of derivatives of potentials (Achenbach, 1973). It is well known that the displacement vector can be written in the form

$$\mathbf{u}^E = \nabla \phi^E + \nabla \times \psi^E \quad (8)$$

Equation (8) satisfies the equations of motion (3) if

$$\nabla \cdot \nabla \phi^E = \frac{\ddot{\phi}^E}{C_L^2}, \quad \nabla \cdot \nabla \psi^E = \frac{\ddot{\psi}^E}{C_T^2} \quad (9)$$

where, if λ and μ denoted Lame's elastic constants, we have

$$C_L^2 = (\lambda + 2\mu)/\rho_0, \quad C_T^2 = \mu/\rho_0 \quad (10)$$

and the additional constraint condition is

$$\nabla \cdot \psi^E = 0. \quad (11)$$

The corresponding elastic stresses are obtained from

$$\mathbf{T}^E = \lambda |\mathbf{S}^E| \mathbf{I} + 2\mu \mathbf{S}^E. \quad (12)$$

Finally, total stresses may be given by

$$\mathbf{T} = \mathbf{T}^T + \mathbf{T}^E. \quad (13)$$

Analysis

The isotropic, homogeneous, elastic spherical shell has uniform temperature $T = 0$, prior to time $t = 0$. The temperature is brought linearly to a uniform temperature rise "T_0" at $t = 0$. The temperature distribution is assumed to have the following form,

$$T = T_0 \{ tH(t) - (t - t_0) H(t - t_0) \} / t_0 \quad (14)$$

where t_0 is a finite heating time and $H(t)$ is the Heaviside step function. The motion of spherical shell subject to spherical symmetric heating is determined from a thermoelastic displacement potential $\phi^T(r, t)$ of equation (4), which satisfies the following equation

$$\frac{\partial^2 \phi^T}{\partial r^2} + \frac{2}{r} \frac{\partial \phi^T}{\partial r} - \frac{1}{c^2} \frac{\partial^2 \phi^T}{\partial t^2} = \frac{1+\nu}{1-\nu} \alpha T \quad (15)$$

where $c = C_L$.

The displacement and the radial and tangential components of the thermal stresses are related to ϕ^T of equations (4) and (6) by

$$\begin{aligned} u^T &= \phi^T_{,r} \\ \sigma_r^T &= -(4\mu/r)\phi^T_{,r} + \rho_0 \phi^T_{,tt} \\ \sigma_\theta^T &= \sigma_\varphi^T = 2\mu(\phi^T_{,r}/r - \nabla^2 \phi^T) + \rho_0 \phi^T_{,tt} \end{aligned} \quad (16)$$

where ∇^2 is given by

$$\nabla^2 = \phi^T_{,rr} + (2/r)\phi^T_{,r}. \quad (17)$$

We shall make use of the Laplace transform on t, defined as

$$\begin{aligned} \bar{\zeta}(r,p) &= \int_0^\infty \zeta(r,t) e^{-pt} dt \\ \zeta(r,t) &= \frac{1}{2\pi i} \int_{\epsilon_0 - i\infty}^{\epsilon_0 + i\infty} \bar{\zeta}(r,p) e^{pt} dp. \end{aligned} \quad (18)$$

Applying the Laplace transform to equation (14), we obtain

$$\bar{T}(r, p)/T_0 = (1 - e^{-pt_0})/(p^2 t_0). \quad (19)$$

The corresponding thermal stresses are

$$\bar{\sigma}_r^T = \bar{\sigma}_\theta^T = -\epsilon T_0 \rho_0 c^2 (1 - e^{-pt_0})/(p^2 t_0) \quad (20)$$

where

$$\epsilon = (1+\nu)\alpha/(1-\nu). \quad (21)$$

These stresses give nonzero boundary tractions on the surfaces of the spherical shell. The effects of removing these stresses are found by solving an ordinary dynamic stress problem, using a displacement potential $\phi^E(r, t)$ which satisfies the following equation (Achenbach, 1973),

$$\frac{\partial^2 \phi^E}{\partial r^2} + \frac{2}{r} \frac{\partial \phi^E}{\partial r} = \frac{1}{c^2} \frac{\partial^2 \phi^E}{\partial t^2}. \quad (22)$$

The boundary conditions of equation (7) are

$$\sigma_r(r,t) = \sigma_r^T(r,t) + \sigma_r^E(r,t) = 0 \quad \text{at } r = a, b. \quad (23)$$

The shell is at rest prior to time $t = 0$ and the initial conditions of displacement $u(r,t)$ are

$$u(r,t) = 0, \ \partial u/\partial t = 0; \ t = 0. \quad (24)$$

The displacement and the radial and tangential components of the stress are related to ϕ^E by

$$\begin{aligned} u^E &= \phi^E_{,r} \\ \sigma_r^E &= -4\mu\phi^E_{,r}/r + \rho_0 \phi^E_{,tt} \\ \sigma_\theta^E &= \sigma_\varphi^E = 2\mu\phi^E_{,r}/r + (\lambda/c^2)\phi^E_{,tt}. \end{aligned} \quad (25)$$

Applying the Laplace transform to equation (22), and solving the transformed equation we obtain

$$\bar{\phi}^E(r,p) = C_1 h_0^{(1)}\left(i\frac{p}{c} r\right) + C_2 h_0^{(2)}\left(i\frac{p}{c} r\right) \quad (26)$$

where C_1 and C_2 are the unknown constants and $h_0^{(1),(2)}(z) = \exp(\pm iz)/(\pm iz)$ are the zeroth-order spherical Hankel functions of the first and second kinds, respectively. The corresponding stress components are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\bar{\sigma}_r^E = C_1\left\{-\frac{4\mu}{r}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}r\right)+\rho_0 p^2 h_0^{(1)}\left(i\frac{p}{c}r\right)\right\}$$
$$+C_2\left\{-\frac{4\mu}{r}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}r\right)+\rho_0 p^2 h_0^{(2)}\left(i\frac{p}{c}r\right)\right\} \quad (27)$$
$$\bar{\sigma}_\theta^E = C_1\left\{\frac{2\mu}{r}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}r\right)+\frac{\lambda p^2}{c^2}h_0^{(1)}\left(i\frac{p}{c}r\right)\right\}$$
$$+C_2\left\{\frac{2\mu}{r}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}r\right)+\frac{\lambda p^2}{c^2}h_0^{(2)}\left(i\frac{p}{c}r\right)\right\}$$

where

$$h_0^{(1),(2)\prime}(z) = dh_0^{(1),(2)}(z)/dz. \quad (28)$$

The unknown constants may be determined from the boundary conditions of equation (23). That is

$$C_1 = (-C_{22}+C_{12})\frac{\bar{p}_0}{\Delta}$$
$$C_2 = (-C_{11}+C_{21})\frac{\bar{p}_0}{\Delta} \quad (29)$$

where

$$\bar{p}_0 = -\frac{\epsilon T_0\rho_0 c^2}{t_0}\frac{(1-e^{-pt_0})}{p^2}$$
$$C_{11} = -\frac{4\mu}{a}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}a\right)+\rho_0 p^2 h_0^{(1)}\left(i\frac{p}{c}a\right)$$
$$C_{12} = -\frac{4\mu}{a}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}a\right)+\rho_0 p^2 h_0^{(2)}\left(i\frac{p}{c}a\right) \quad (30)$$
$$C_{21} = -\frac{4\mu}{b}\left(i\frac{p}{c}\right)h_0^{(1)\prime}\left(i\frac{p}{c}b\right)+\rho_0 p^2 h_0^{(1)}\left(i\frac{p}{c}b\right)$$
$$C_{22} = -\frac{4\mu}{b}\left(i\frac{p}{c}\right)h_0^{(2)\prime}\left(i\frac{p}{c}b\right)+\rho_0 p^2 h_0^{(2)}\left(i\frac{p}{c}b\right)$$
$$\Delta = C_{11}C_{22}-C_{12}C_{12}.$$

The Ray Integrals

Substituting equations (29) into (26) we obtain

$$\bar{\phi}^E = -\frac{\bar{p}_0}{\Delta}C_{22}\left\{h_0^{(1)}\left(i\frac{p}{c}r\right)-\frac{C_{21}}{C_{22}}h_0^{(2)}\left(i\frac{p}{c}r\right)\right\}$$
$$-\frac{\bar{p}_0}{\Delta}C_{11}\left\{h_0^{(2)}\left(i\frac{p}{c}r\right)-\frac{C_{12}}{C_{11}}h_0^{(1)}\left(i\frac{p}{c}r\right)\right\}. \quad (31)$$

Here, we introduce two reflection coefficients R and Q, which are defined as

$$R = -C_{12}/C_{22}, \quad Q = -C_{21}/C_{11} \quad (32)$$

where R denotes the reflection coefficient for an outgoing spherical wave and Q denotes that for an incoming wave. The denominator in equations (29) is written as

$$\Delta = C_{11}C_{22}(1-RQ). \quad (33)$$

The inverse of the denominator can be expanded into a series of RQ as follows:

$$\Delta^{-1} = (C_{11}C_{22})^{-1}(1+RQ+R^2Q^2+\text{-----}) \ (|RQ|<1). \quad (34)$$

The series in equation (34) is convergent when the real part of p is positive. Substituting equations (34) into (31), we obtain

$$\bar{\phi}^E(r,p) = \sum_{j=0}^{\infty}\{\bar{\phi}_j(r,p)+\bar{\psi}_j(r,p)\} \quad (35)$$

where

$$\left.\begin{aligned}
\bar{\phi}_0(r,p) &= -\frac{\bar{p}_0}{C_{11}}h_0^{(1)}\left(i\frac{p}{c}r\right)\\
\bar{\phi}_1(r,p) &= -R(p)\frac{\bar{p}_0}{C_{11}}h_0^{(2)}\left(i\frac{p}{c}r\right)\\
\bar{\phi}_j(r,p) &= R(p)Q(p)\bar{\phi}_{j-2}(r,p) \qquad (j\geqq 2)
\end{aligned}\right\} \quad (36)$$

$$\left.\begin{aligned}
\bar{\psi}_0(r,p) &= -\frac{\bar{p}_0}{C_{22}}h_0^{(2)}\left(i\frac{p}{c}r\right)\\
\bar{\psi}_1(r,p) &= -Q(p)\frac{\bar{p}_0}{C_{22}}h_0^{(1)}\left(i\frac{p}{c}r\right)\\
\bar{\psi}_j(r,p) &= R(p)Q(p)\bar{\psi}_{j-2}(r,p) \qquad (j\geqq 2)
\end{aligned}\right\} \quad (37)$$

Finally, the complete solution of displacement potential is given by

$$\phi^E(r,t) = \frac{1}{2\pi i}\sum_{j=0}^{\infty}\int_{\epsilon_0-i\infty}^{\epsilon_0+i\infty}\{\bar{\phi}_j(r,p)+\bar{\psi}_j(r,p)\}e^{pt}dp. \quad (38)$$

Each term of the series is called a ray integral. The inversion of Laplace transform of the functions $\bar{\phi}_j$ and $\bar{\psi}_j$ is obtained as follows: First applying the inversion of Laplace transform of equation (18) to the functions $\bar{\phi}_0$, $\bar{\phi}_1$, $\bar{\psi}_0$, and $\bar{\psi}_1$, we obtain these functions in closed form analytically. Second, the higher-order terms of ϕ_j and ψ_j are obtained through the convolution integral. That is

$$\phi_j(r,t) = q(t)*\phi_{j-2}(r,t) = \int_0^t q(\tau)\phi_{j-2}(r,t-\tau)d\tau$$
$$\psi_j(r,t) = q(t)*\psi_{j-2}(r,t) = \int_0^t q(\tau)\psi_{j-2}(r,t-\tau)d\tau \quad (39)$$

where

$$q(t) = \frac{1}{2\pi i}\int_{\epsilon_0-i\infty}^{\epsilon_0+i\infty}R(p)q(p)e^{pt}dp$$
$$\phi_{j-2}(r,t) = \frac{1}{2\pi i}\int_{\epsilon_0-i\infty}^{\epsilon_0+i\infty}\bar{\phi}_{j-2}(r,p)e^{pt}dp \quad (40)$$
$$\psi_{j-2}(r,t) = \frac{1}{2\pi i}\int_{\epsilon_0-i\infty}^{\epsilon_0+i\infty}\bar{\psi}_{j-2}(r,p)e^{pt}dp$$

Substituting equation (30) into equation (36), we can calculate the inverse Laplace transform of function $\bar{\phi}_0$ and $\bar{\phi}_1$ as follows:

$$\phi_0(r,t) = \frac{\epsilon T_0 a^3}{r}\sum_{i=0}^{2}\frac{e^{\alpha_i(t+a/c-r/c)}H\left(t+\frac{a}{c}-\frac{r}{c}\right)}{3\left(\frac{\alpha_i a}{c}\right)^2+4n_0\left(\frac{\alpha_i a}{c}\right)+4\kappa} \quad (41)$$

$$\phi_1(r,t) = -\frac{\epsilon T_0 a^3}{r}\sum_{i=0}^{4}$$
$$\frac{\left(\frac{\alpha_i b}{c}+\gamma\right)\left(\frac{\alpha_i b}{c}+\gamma^*\right)e^{\alpha_i(t+a/c+r/c-2b/c)}H\left(t+\frac{a}{c}+\frac{r}{c}-\frac{2b}{c}\right)}{5\left(\frac{\alpha_i}{c}\right)^4a^2b^2+8n_0\left(\frac{\alpha_i}{c}\right)^3ab(b-a)+12\left(\frac{\alpha_i}{c}\right)^2(\kappa a+\kappa b^2-n_0^2ab)+16n_0\kappa\left(\frac{\alpha_i}{c}\right)(a-b)+16\kappa^2}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where

$$\kappa=\mu/(\lambda+2\mu),\ n_0=2\kappa,\ \alpha_0=0,\ \alpha_1=-\frac{c}{a}\gamma,$$

$$\alpha_2=-\frac{c}{a}\gamma^*,\ \alpha_3=\frac{c}{b}\gamma,\ \alpha_4=\frac{c}{b}\gamma^* \tag{42}$$

and γ and γ^* are denoted as

$$\gamma=2\kappa\left(1-\sqrt{1-1/\kappa}\right),\ \gamma^*=2\kappa\left(1+\sqrt{1-1/\kappa}\right). \tag{43}$$

The function $q(t)$ of equation (40) is given by

$$q(t)=\delta(t+2a/c-2b/c)+q_c(t) \tag{44}$$

where

$$q_c(t)=\sum_{i=1}^{4}\frac{2n_0\left(\frac{a}{c}-\frac{b}{c}\right)\left(\frac{\alpha_i^3ab}{c^2}-4\kappa\alpha_i\right)e^{\alpha_i(t-2b/c+2a/c)}H\left(t+2\frac{a}{c}-2\frac{b}{c}\right)}{2\left(\frac{\alpha_i^3a^2b^2}{c^4}\right)+3n_0\left(\frac{\alpha_i^2ab}{c^3}\right)(b-a)+4\frac{\alpha_i}{c^2}(\kappa a^2+\kappa b^2-n_0ab)+\frac{4n_0\kappa}{c}(a-b)}. \tag{45}$$

Substituting equations (45) into (39), the function $\phi_n(n\geqq 2)$ is obtained as follows:

$$\phi_{2n}(r,t)=\phi_{2n-2}\left(r,t+\frac{2a}{c}-\frac{2b}{c}\right)+\int_{2(b-a)/c}^{t+(a-r)/c+2(n-1)(a-b)/c}q_c(\tau)\phi_{2n-2}(r,t-\tau)d\tau H\left[t+\frac{a}{c}-\frac{r}{c}+2n\left(\frac{a}{c}-\frac{b}{c}\right)\right] \tag{46}$$

$$\phi_{2n+1}(r,t)=\phi_{2n-1}\left(r,t+\frac{2a}{c}-\frac{2b}{c}\right)+\int_{2(b-a)/c}^{t+(a+r-2b)/c+2(n-1)(a-b)/c}q_c(\tau)\phi_{2n-1}(r,t-\tau)d\tau H\left[t+\frac{a}{c}+\frac{r}{c}-2\frac{b}{c}+2n\left(\frac{a}{c}-\frac{b}{c}\right)\right]. \tag{47}$$

Similarly, the inversion of Laplace transform of the functions $\overline{\psi_0}$ and $\overline{\psi_1}$ is obtained from equations (37) as

$$\psi_0(r,t)=\frac{\epsilon T_0b^3}{r}\sum_{j=0}^{2}\frac{e^{\alpha_j(t+r/c-b/c)}H\left[t+\frac{r}{c}-\frac{b}{c}\right]}{3\left(\frac{\alpha_jb}{c}\right)^2-4n_0\left(\frac{\alpha_jb}{c}\right)+4\kappa}$$

$$\psi_1(r,t)=-\frac{\epsilon T_0b^3}{r}\sum_{j=0}^{4}\frac{\left(\frac{\alpha_ja}{c}-\gamma\right)\left(\frac{\alpha_ja}{c}-\gamma^*\right)}{g(\alpha_j)+\alpha_jg'(\alpha_j)}\times e^{\alpha_j(t+2a/c-b/c-r/c)}H\left[t+2\frac{a}{c}-\frac{b}{c}-\frac{r}{c}\right] \tag{48}$$

where

$$\alpha_0=0,\ \alpha_1=\gamma(c/b),\ \alpha_2=\gamma^*(c/b),\ \alpha_3=-\gamma(c/a),\ \alpha_4=-\gamma^*(c/a) \tag{49}$$

and $g(\alpha_j)$ is denoted as

$$g(\alpha_j)=\frac{\alpha_j^4a^2b^2}{c^4}+2n_0\frac{\alpha_j^3ab}{c^3}(b-a)+2n_0\frac{\alpha_j^2}{c^2}(a^2-4\kappa ab+b^2)+4n_0^2\frac{\alpha_j}{c}(a-b)+4n_0^2. \tag{50}$$

Substituting $q(t)$ of equations (44) into (39), the function $\psi_n\ (n\geqq 2)$ is obtained as follows:

$$\psi_{2n}(r,t)=\psi_{2n-2}\left(r,t+\frac{2a}{c}-\frac{2b}{c}\right)+\int_{2(b-a)/c}^{t+(r-b)/c+2(n-1)(a-b)/c}q_c(\tau)\psi_{2n-2}(r,t-\tau)d\tau H\left[t+\frac{r}{c}-\frac{b}{c}+2n\left(\frac{a}{c}-\frac{b}{c}\right)\right] \tag{51}$$

$$\psi_{2n+1}(r,t)=\psi_{2n-1}\left(r,t+\frac{2a}{c}-\frac{2b}{c}\right)+\int_{2(b-a)/c}^{t+(2a-r-b)/c+2(n-1)(a-b)/c}q_c(\tau)\psi_{2n-1}(r,t-\tau)d\tau H\left[t-\frac{r}{c}-\frac{b}{c}+\frac{2a}{c}+2n\left(\frac{a}{c}-\frac{b}{c}\right)\right]. \tag{52}$$

Therefore, the elastic displacement potential ϕ^E is given from equation (38) by

$$\phi^E(r,t)=\sum_{j=0}^{\infty}\{\phi_j(r,t)+\psi_j(r,t)\}. \tag{53}$$

The corresponding displacement and stresses are

$$u^E(r,t)=\phi^E,r$$

$$\sigma_r^E(r,t)=\rho_0c^2\left[\phi^E(r,t),_{rr}+\frac{2\nu}{1-\nu}\frac{1}{r}\phi^E(r,t),_r\right] \tag{54}$$

$$\sigma_\theta^E(r,t)=\rho_0c^2\left[\frac{\nu}{1-\nu}\phi^E(r,t),_{rr}+\frac{1}{(1-\nu)r}\phi^E(r,t),_r\right].$$

Finally, the complete solution of the problem is obtained by the summation of the solutions of thermal and dynamic problems as follows:

$$u=u^E+u^T,\ \sigma_r=\sigma_r^E+\sigma_r^T,\ \sigma_\theta=\sigma_\theta^E+\sigma_\theta^T. \tag{55}$$

Numerical Results and Discussion

To show the interfering effect between the inward and the outward waves in a sphere subjected suddenly to a uniform temperature rise throughout the sphere, we performed numerical calculations for two isotropic spherical shells with $b/a = 2$ and 5. The temperature field is given by $T=T_0H(t)$. The Poisson's ratio is 0.3. The numerical integrations were carried out using Simpson's formula.

The results of the numerical evaluation of stress variation are illustrated in Figs. 1 to 5. In these figures we used the nondimensional variables $t^* = ct/a$ and $\sigma_i^* = \sigma_i/(\epsilon T_0\rho_0c^2)$.

The time history of tangential stress for a spherical shell with $b/a = 2$ at various locations is shown in Fig. 1. Figure 2 shows that in a spherical shell $b/a=5$.

In Fig. 1, as the stress $\sigma_\theta^* = -H(t^*)$ is suddenly applied to the entire sphere, the stress waves occur at the internal and the external surfaces at the same time. The wave at the internal surface proceeds radially outward, while the wave at the external surface proceeds inward. Therefore, the stress variation curves at the internal and external surfaces have stress discontinuities at $t^* = 1, 2, \ldots\ldots, 10$. The maximum peak stress, which occurs at r/a when $t^*=9$, is about $\sigma_\theta^* = 1.8$.

In Fig. 2, we observe that the time history of dynamic stresses in a spherical shell with $b/a = 5$ is similar to that in Fig. 1 as a result of multiple reflection of waves. In this case, the maximum peak stress, which occurs at $r=a$ when $t^*=5$, is about $\sigma_\theta^* = 4.8$. It is remarkable that this value is about 2.7 times higher than that in Fig. 1.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

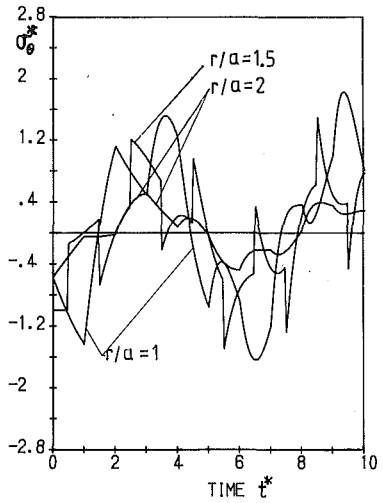


Fig. 1 Variation of tangential stress in the case of $b/a = 2$

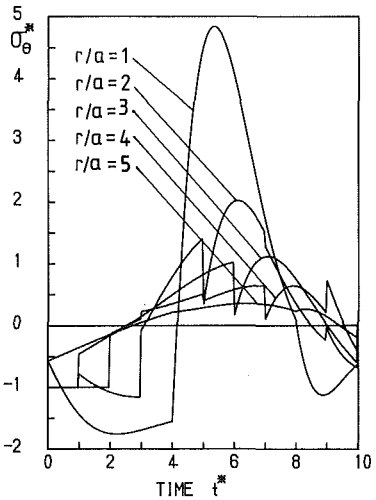


Fig. 2 Variation of tangential stress in the case of $b/a = 5$

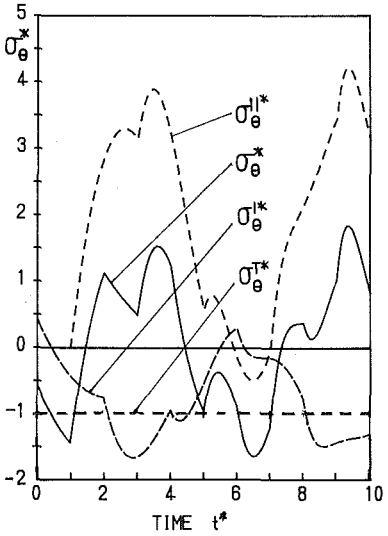


Fig. 3 Time historical relation of tangential stresses σ_θ^{1*}, σ_θ^{11*}, and σ_θ^{T*} to σ_θ^* at $r/a = 1$ in the case of $b/a = 2$

In order to clarify the reason, we investigate the interfering effects between the inward stress waves and the outward stress waves in a sphere subjected to a suddenly uniform heating.

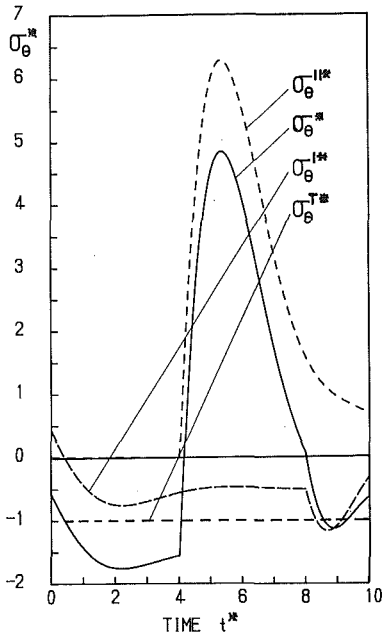


Fig. 4 Time historical relation of tangential stresses σ_θ^{1*}, σ_θ^{11*}, and σ_θ^{T*} to σ_θ^* at $r/a =$ in the case of $b/a = 5$

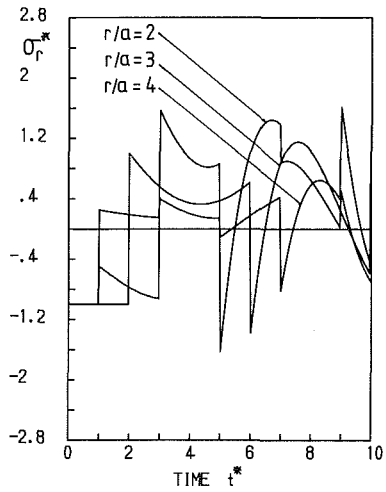


Fig. 5 Variation of radial stress in the case of $b/a = 5$

Figure 3 shows the time historical relation of tangential stresses σ_θ^{1*}, σ_θ^{11*}, and σ_θ^{T*} to σ_θ^* at the internal surface of a spherical shell with $b/a = 2$. Figure 4 shows the time historical relation of tangential stresses in a spherical shell with $b/a = 5$. In these figures, the nondimensional stresses σ_θ^{1*}, σ_θ^{11*}, and σ_θ^{T*} denote

$$\sigma_\theta^{1*} = \sum_{j=0}^{\infty}\left[\phi_j(r,t)_{,rr} + \frac{2\nu}{(1-\nu)}\frac{1}{r}\phi_j(r,t)_{,r}\right]/\epsilon T_0$$

$$\sigma_\theta^{11*} = \sum_{j=0}^{\infty}\left[\psi_j(r,t)_{,rr} + \frac{2\nu}{(1-\nu)}\frac{1}{r}\psi_j(r,t)_{,r}\right]/\epsilon T_0 \tag{56}$$

$$\sigma_\theta^{T*} = -H(t^*)$$

and $\sigma_\theta^* = \sigma_\theta^{1*} + \sigma_\theta^{11*} + \sigma_\theta^{T*}$.

In equation (56) σ_θ^{1*} is the tangential stress caused by the wave created at the internal surface, while σ_θ^{11*} is that at the external surface. It is seen from Fig. 3 that the value of σ_θ^{11*} at $t^* = 3.4$ is 3.8, while that of σ_θ^{1*} is -1.5.

Considering $\sigma_\theta^{T*} = -1$, the total tangential stress σ_θ^* is about

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

1.3. In the case of $b/a = 5$ in Fig. 4, at $t^* = 5.3$, the value of σ_θ^{1*} is about -0.5, while σ_θ^{11*} is about 6.3. Considering $\sigma_\theta^{T*} = -1$, the total tangential stress σ_θ^* is 4.8.

Comparing these two cases, we observe that the maximum peak value of tangential stress σ_θ^{1*} at the internal surface in the case of $b/a = 5$ is almost the same order of the stress in the case of $b/a = 2$, but the peak value of tangential stress σ_θ^{11*} in the case of $b/a = 5$ is remarkably higher than that in the case of $b/a = 2$. Therefore, it can be assumed that, as the ratio of b/a is increasing, the tangential stress σ_θ^{11*} at the internal surface becomes higher. The peak value of tangential stress σ_θ^* at the internal surface becomes higher and higher as the result of increasing the tangential stress σ_θ^{11*}. Since the solid sphere is the limit case, where the ratio of b/a is infinity, it may be assumed that the remarkably high stress occurs at the center of the solid sphere. The same phenomenon observed in a cylindrical rod introduced by Ho (1976) is called Stress-Focusing Effect.

Figure 5 shows the time history of radial stresses for a spherical shell with thickness $b/a = 5$. It is seen that, compared to the tangential stress, the radial stress in a spherical shell is insignificant.

Conclusion

This paper is concerned with a method for calculating the dynamic thermoelastic stress distribution in a hollow sphere. Adopting the Goodier's concept, the dynamic thermoelastic problem in a sphere is decomposed into the particular form of stress problem, which is solved by using the thermoelastic potential, and the homogeneous form of dynamic stress problems, which is solved by using stress functions. Applying the ray theory, we obtained the general exact solution for transient waves induced by sudden heating. When a hollow sphere is subjected suddenly to a uniform temperature rise throughout the shell, stress waves occur at the internal and external surfaces the moment thermal impact is applied. The wave at the internal surface proceeds radially outward, while the wave at the external surface proceeds inward. During instantaneous heating, the interfering effects of these waves can cause a very high dynamic stress in a sphere. The numerical results show that, as the ratio of b/a increases, the tangential stress on the internal surface becomes higher and higher because of the increase in the tangential stress σ_θ^{11*} created at the external surface.

Acknowledgments

The author is especially grateful to Professor Emeritus A. Atsumi of Tohoku University and Prof. Y. H. Pao of Cornell University for their invaluable suggestions. This work was partly supported by Grant-Aid of the National Science Foundation from the Ministry of Education in Japan.

References

Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North-Holland Publishing Company, Amsterdam.

Goodier, J. N., 1937, "Integration of Thermoelastic Equations," *Phil. Mag.*, Vol. 37, pp. 1017.

Ho, C. H., 1976, "Stress-Focusing Effect in a Uniformly Heated Cylindrical Rod," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 404–468.

Pao, Y. H., and Ceranoglue, A. N., 1978, "Determination of Transient Response of a Thick-Walled Spherical Shell by Ray Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 114–122.

Sternberg, E., and Chakravorty, J. G., 1959, "Thermal Shock in an Elastic Body with a Spherical Cavity," *Quart. Appl. Math.*, Vol. 17, pp. 205–218.

Tsui, T., and Kraus, H., 1965, "Thermal Stress-Wave Propagation in Hollow Elastic Spheres," *J. Acoust. Soc. Am.*, Vol. 37, No. 4, pp. 730–737.

Zaker, T. A., 1966, "Dynamic Thermal Shock in a Spherical Elastic Shell of Arbitrary Thickness," *Proc. 5th U.S. Nat. Cong. Appl. Mech.*, p. 183.

Zaker, T. A., 1968, "Dynamic Thermal Shock in a Hollow Sphere," *Quart. Appl. Math.*, Vol. 26, p. 503.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Creasing Singularities in Compressible Elastic Materials

S. A. Silling
Division of Engineering,
Brown University,
Providence, RI 02912
Assoc. Mem. ASME

The deformation and stress fields near a crease in the surface of an isotropic, compressible elastic body in equilibrium are in general singular. The order of the singularity depends on the constitutive behavior in the limit of zero volume. In this paper the problem of the creasing of a semicircular cylinder folded about its axis is solved exactly for a family of materials that includes the Blatz-Ko material (a foam rubber). Conditions under which strong ellipticity is lost are derived.

Introduction

In finite elastostatics, stress singularities may occur as a result of gross changes in shape induced by a deformation. A common example of such a singularity is found in the creasing or folding of a body along an initially flat surface. The purpose of this paper is to discuss the nature of this singularity in a family of compressible, isotropic elastic materials.

1 Creasing of a Semicircular Region

Let the undeformed body be a stress-free homogeneous cylinder of semi-circular cross-section of radius r_0 (see Fig. 1). Plane strain is assumed. Let $\mathbf{x}$ be the position vector in the cross-sectional plane of a material particle in the undeformed configuration, with polar coordinates (r, θ) as shown in the figure. Let $\mathbf{y} = \hat{\mathbf{y}}(\mathbf{x})$ be the image of $\mathbf{x}$ in the deformed configuration, with polar coordinates given by

$$R = \hat{R}(r,\theta),\ \phi = \hat{\phi}(r,\theta),\ 0 \le r \le r_0,\ -\pi/2 \le \theta \le \pi/2 \qquad (1)$$

where $\hat{R}$ and $\hat{\phi}$ are functions that are assumed to be twice continuously differentiable, except possibly at $r = 0$. Let $\mathbf{F}$ denote the deformation gradient tensor field defined on the undeformed region, $\mathbf{F} = \nabla\hat{\mathbf{y}}$.

Let β be a constant, $1 \le \beta \le 2$. It is assumed that $\theta = \pm\pi/2$ is mapped to $\phi = \pm\beta\pi$, and that $r = 0$ is mapped to $R = 0$. The surfaces $\phi = \pm\pi$ are assumed to sustain zero shear traction. A crease corresponds to $\beta = 2$, while the undeformed condition corresponds to $\beta = 1$. The outer circular boundary is assumed to be subjected to a constant normal compressive traction p_0 in the deformed configuration and zero shear traction.

For isotropic *incompressible* materials, the deformation which solves this problem is a special case of the class of the universal deformations investigated by Singh and Pipkin (1965) which map wedges into wedges:

$$\hat{R}(r,\theta) = r/\sqrt{\beta},\ \hat{\phi}(r,\theta) = \beta\theta. \qquad (2)$$

Let τ be the Cauchy stress tensor, with in-plane components in the polar coordinate system denoted τ_{RR}, $\tau_{\phi\phi}$, $\tau_{R\phi}$. Because of the symmetries in the deformation described by equation (2), $\tau_{R\phi} \equiv 0$, while τ_{RR} and $\tau_{\phi\phi}$ are independent of ϕ. The equilibrium equation for forces in the tangential direction is automatically satisfied because of these symmetries. The equilibrium equation for radial forces simplifies to

$$\frac{d\tau_{RR}}{dR} + \frac{\tau_{RR} - \tau_{\phi\phi}}{R} = 0,\ 0 < R \le R_0 \qquad (3)$$

with boundary condition

$$\tau_{RR}(R_0) = -p_0, \qquad (4)$$

where $R_0 = r_0/\sqrt{\beta}$. The hydrostatic pressure field may be found for any isotropic incompressible material by integrating equation (3).

For three special types of isotropic *compressible* materials, including harmonic materials, Carroll (1988) found several classes of universal deformations, one of which provides solutions to this creasing problem. However, this assumed set of constitutive relations is inflexible in the types of behavior it can model in the limit of zero volume; hence the motivation to study creasing singularities in alternative materials.

For the class of isotropic compressible elastic materials that will be specified, we will attempt to find solutions in which R

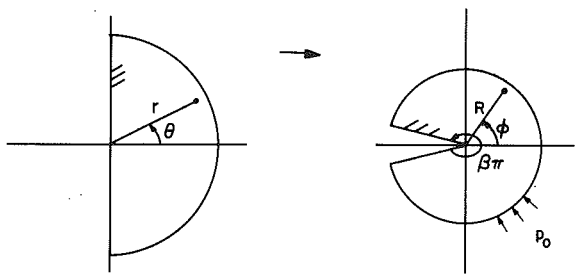


Fig. 1 Undeformed and deformed configurations

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 10, 1989; final revision, Dec. 18, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

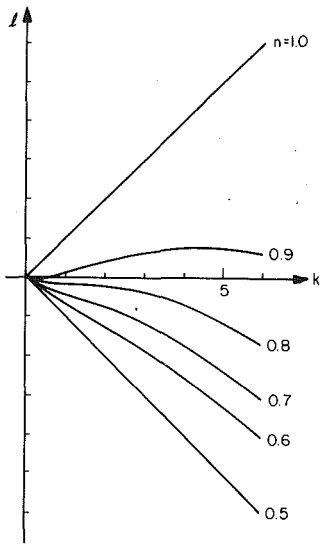


Fig. 2 Contours of constant n

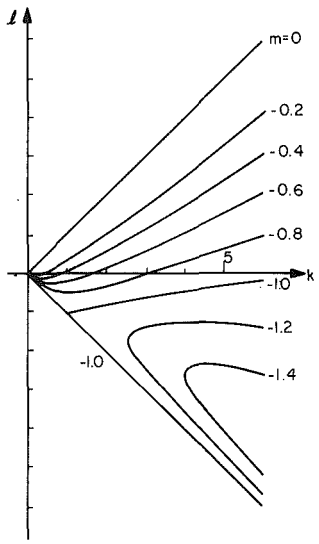


Fig. 3 Contours of constant m

is independent of θ and in which $\phi = \beta\theta$. Under these assumptions, equations (3) and (4) continue to hold, where R_0 is the outer radius of the deformed body. Equilibrium of forces in the tangential direction is again satisfied automatically. It is easiest to formulate the remaining equations in the deformed configuration. We assume that there is a function $\hat{r}(\cdot)$ on $[0, R_0]$ such that the inverse mapping of $\hat{\mathbf{y}}$ is given by

$$r = \hat{r}(R),\ \theta = \phi/\beta,\ -\beta\pi/2 \le \phi \le \beta\pi/2,\ 0 \le R \le R_0. \quad (5)$$

$\hat{r}$ is assumed to be twice continuously differentiable except possibly at $R = 0$.

Let $W(\mathbf{F})$ be the strain energy density function for plane strain. We assume $\det\mathbf{F} > 0$. Let $\mathbf{F} = \mathbf{VQ}$ be the two-dimensional left polar decomposition of $\mathbf{F}$, so that $\mathbf{V}$ is a symmetric positive definite tensor, $\mathbf{Q}$ is a proper orthogonal tensor, and $\mathbf{V} = \sqrt{\mathbf{FF}^T}$. Let $\mathbf{e}_1$, $\mathbf{e}_2$ be a principal basis for $\mathbf{V}$, so that the components of $\mathbf{V}$ in this basis are

$$[\mathbf{V}]^e = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad (6)$$

where λ_1, λ_2 are the principal stretches of $\mathbf{F}$. In plane strain, for an isotropic compressible elastic material, there is a function $\tilde{W}(\cdot,\cdot)$ such that $W(\mathbf{F}) = \tilde{W}(\lambda_1, \lambda_2)$. Isotropy further requires that $\tilde{W}(\lambda_1, \lambda_2) = \tilde{W}(\lambda_2, \lambda_1)$ for any λ_1, λ_2. The in-plane principal Cauchy stresses are given by

$$\tau_\alpha = \frac{\lambda_\alpha}{\lambda_1\lambda_2}\frac{\partial\tilde{W}}{\partial\lambda_\alpha}\ (\text{no sum}), \quad (7)$$

where $\alpha = 1, 2$. Under the assumptions in equation (5), at an arbitrary point $\mathbf{x} \neq \mathbf{0}$ in the body the principal basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ coincide with the radial and tangential directions at $\hat{\mathbf{y}}(\mathbf{x})$. The principal stretches are given as a function of deformed radius by

$$\lambda_1(R) = \lambda_R(R) = \frac{1}{\hat{r}'(R)},\ \lambda_2(R) = \lambda_\phi(R) = \frac{\beta R}{\hat{r}(R)},\ 0 < R \le R_0 \quad (8)$$

where $\hat{r}' = d\hat{r}/dR$. As in the incompressible case, equation (5) implies that $\tau_{R\phi} \equiv 0$ and that τ_{RR} and $\tau_{\phi\phi}$ are independent of ϕ. It follows that

$$\tau_{RR}(R) = \frac{\hat{r}(R)}{\beta R}\frac{\partial\tilde{W}}{\partial\lambda_1}\left(\frac{1}{\hat{r}'(R)}, \frac{\beta R}{\hat{r}(R)}\right),$$

$$\tau_{\phi\phi}(R) = \hat{r}'(R)\frac{\partial\tilde{W}}{\partial\lambda_2}\left(\frac{1}{\hat{r}'(R)}, \frac{\beta R}{\hat{r}(R)}\right). \quad (9)$$

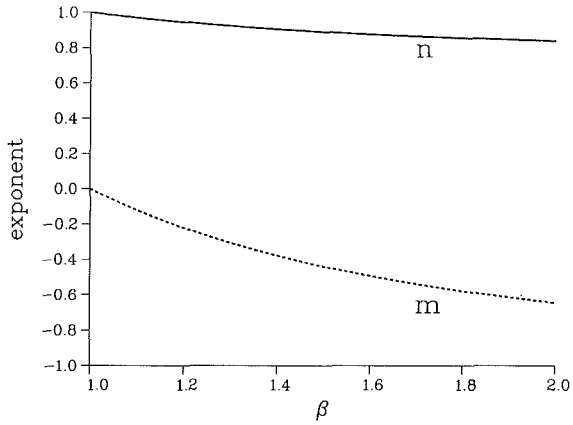


Fig. 4 n and m as a function of β for the Blatz-Ko material

The result of substituting equations (9) into (3) is a nonlinear second-order ordinary differential equation for $\hat{r}$ as a function of R. The appropriate boundary conditions are

$$\hat{r}(0) = 0,\ \tau_{RR}(R_0) = -p_0. \quad (10)$$

The family of isotropic materials to be considered here has a strain energy density function (in plane strain) given by

$$\tilde{W}(\lambda_1,\lambda_2) = a(\lambda_1^{-k}\lambda_2^{-l} + \lambda_2^{-k}\lambda_1^{-l} - 2) + a(k+l)(\lambda_1\lambda_2 - 1), \quad (11)$$

where k, l, and a are constants, $a > 0$. We assume

$$k \ge l,\ k + l > 0. \quad (12)$$

This first of equation (12) is assumed for the sake of definiteness, although k and l are essentially interchangeable. The second is assumed in order to enforce the physically motivated requirement that τ_1 and τ_2 be compressive in a process of isotropic crushing, i.e., $\lambda_1 = \lambda_2 \to 0$.

The principal stresses are found from equations (7) and (11) to be

$$\tau_1 = -a(k\lambda_1^{-k-1}\lambda_2^{-l-1} + l\lambda_2^{-k-1}\lambda_1^{-l-1}) + a(k+l),$$

$$\tau_2 = -a(l\lambda_1^{-k-1}\lambda_2^{-l-1} + k\lambda_2^{-k-1}\lambda_1^{-l-1}) + a(k+l). \quad (13)$$

If $k \neq l$, then the material satisfies the plane-strain Baker-Ericksen inequality, i.e., $(\tau_1 - \tau_2)(\lambda_1 - \lambda_2) > 0$ whenever $\lambda_1 \neq \lambda_2$.

If $l < 0$, and if λ_2 is held constant while $\lambda_1 \to 0$, equations (13) show that $\tau_1 \to -\infty$ while $\tau_2 \to +\infty$. Thus, the small-volume behavior of the materials with $l < 0$ is somewhat improbable physically.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

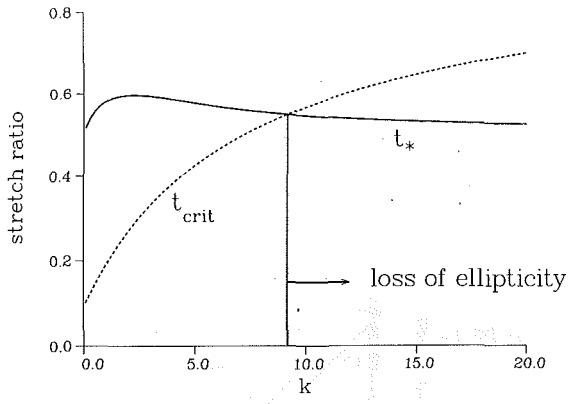


Fig. 5 Ratio of radial to tangential stretch for loss of ellipticity and for the creasing problem, $l=0$

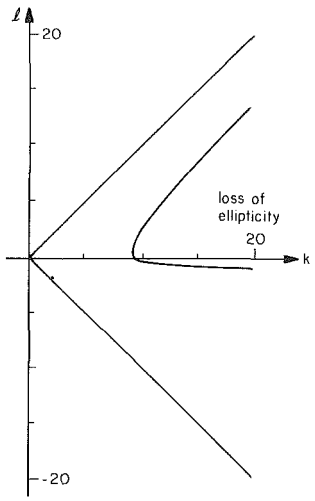


Fig. 6 Region in the (k, l)-plane in which strong ellipticity is lost in the creasing problem

In what follows, we will be interested in how the stress singularity depends on k and l, since these constants determine the constitutive behavior of the material upon crushing ($\lambda_1 \to 0$, $\lambda_2 \to 0$). The term involving $\lambda_1\lambda_2$ in equation (11), which determines the large-stretch response of the material, has no effect on the singularity. A wide variety of responses in the limit of zero volume may be represented by the family of materials in equation (11) for appropriate choices of k and l. Of particular interest is the foam rubber material studied by Blatz and Ko (1962), for which $k=2$ and $l=0$. Note that if $k=l$, the material in equation (11) is fluidlike, since W is then a function of the Jacobian determinant $J = \lambda_1\lambda_2$ only.

To solve the two-point boundary value problem in equations (3), (10), assume a solution of the form

$$\hat{r}(R) = CR^n \tag{14}$$

where C and n are positive constants. The associated principal stretches are

$$\lambda_1 = \lambda_R = \frac{R^{1-n}}{nC},\ \lambda_2 = \lambda_\phi = \frac{\beta R^{1-n}}{C}. \tag{15}$$

Substituting equation (15) into equations (9) and (3) confirms that this assumed form is a solution of equation (3) for any positive value of C provided

$$[(n-1)(k+l+2)+1][k(\beta n)^k + l(\beta n)^l] - [k(\beta n)^l + l(\beta n)^k] = 0. \tag{16}$$

For $1 \le \beta \le 2$, and for $k+l>0$, the algebraic equation (16) has exactly one root n in the interval

$$1/\beta \le n \le 1. \tag{17}$$

This root has the value $n=1$ for $\beta=1$ and decreases monotonically with increasing β. Roots of equation (16) less than $1/\beta$ exist for certain values of k and l, but these all lead to infinite total strain energy and may therefore be disregarded. The stresses associated with equation (14) are

$$\tau_{RR} = a(k+l) - a\left(\frac{n}{\beta^{k+l+1}}\right) C^{k+l+2}[k(\beta n)^k + l(\beta n)^l]R^m,$$

$$\tau_{\phi\phi} = a(k+l) - a\left(\frac{n}{\beta^{k+l+1}}\right) C^{k+l+2}[l(\beta n)^k + k(\beta n)^l]R^m, \tag{18}$$

$$\tau_{R\phi} = 0$$

where

$$m = (n-1)(k+l+2). \tag{19}$$

Combining equations (16), (17), and (19) shows that

$$-2 < m \le 0. \tag{20}$$

For $\beta > 1$, the case $m=0$ (bounded stresses) occurs only in the fluidlike material, $k=l$. For other materials, Figs. 2 and 3 show contours of constant n and m as a function of k and l for the case of creasing, $\beta=2$. Figure 4 shows n and m as a function of β for the Blatz-Ko material ($k=2$, $l=0$).

The second of equation (10) and the first of equation (18) imply

$$C = r_0^{1-n}\left[\frac{\beta^{1+k+l}(p_0 + a(k+l))}{na(k(\beta n)^k + l(\beta n)^l)}\right]^{n/(k+l+2)}. \tag{21}$$

The denominator in equation (21) is always positive under the present assumptions, so a suitable C can be found whenever $p_0 > -a(k+l)$. This restriction is not troublesome because equation (13) shows that the material cannot sustain plane-strain isotropic extension such that $\tau_1 = \tau_2 \ge a(k+l)$.

Computation of the strain energy density using equations (11), (14), and (15) shows that

$$W \sim r^{(n-1)(k+l)/n} \text{ as } r \to 0. \tag{22}$$

Thus, the total strain energy near the singularity is finite provided

$$(n-1)(k+l)/n + 2 > 0, \tag{23}$$

which comes from evaluation of the area-integral of W in the undeformed configuration. The inequality (23) is always satisfied under the present assumptions, as may be confirmed by combining equations (19) and (20). So the singularity contains finite total strain energy.

From now on we assume $\beta=2$ (creasing). In order for the assumption that the surfaces $\theta = \pm\pi/2$ map onto $\phi = \pm\pi$ to hold in a model of creasing, τ_{RR} must be nonpositive along these surfaces. Combining equations (16), (18), (19), and (20) shows that this condition is met if and only if

$$m > -1 \text{ and } p_0 \ge \frac{a(k+l)\,|m|}{m+1} = p_{\min}(k,l) \tag{24}$$

where $p_{\min}(k,l)$ is defined in equation (24). The combinations of k and l for which $m \le -1$ may be seen in Fig. 3. All such materials have $l<0$, and the previous discussion showed that these materials are physically improbable. The second of equation (24) shows that if the first of equation (24) holds, then there is some p_0 such that creasing occurs. Note that the necessary confining pressure becomes unboundedly large as $m \to -1$ from above.

For the materials in which $m < -1$, there are still other strange aspects of the solution. Assume that contact at the crease is maintained, and the solution discussed above there-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

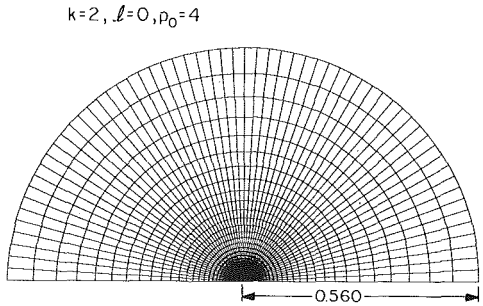


Fig. 7 Deformed mesh for the Blatz-Ko material, $p > p_{min}$

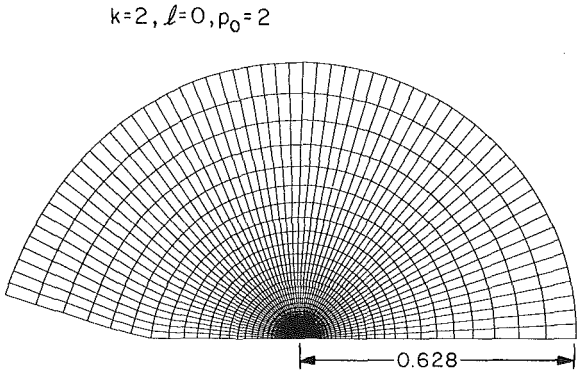


Fig. 8 Deformed mesh for the Blatz-Ko material, $p < p_{min}$

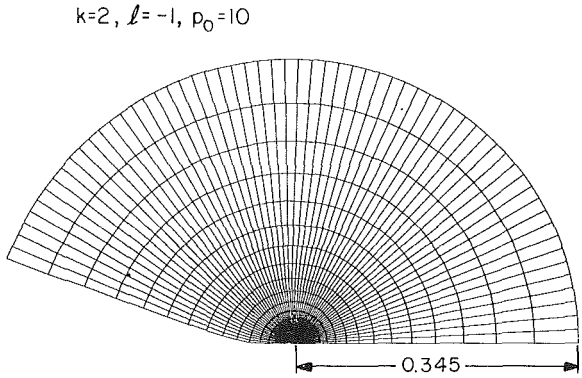


Fig. 9 Deformed mesh for a material with $m < -1$

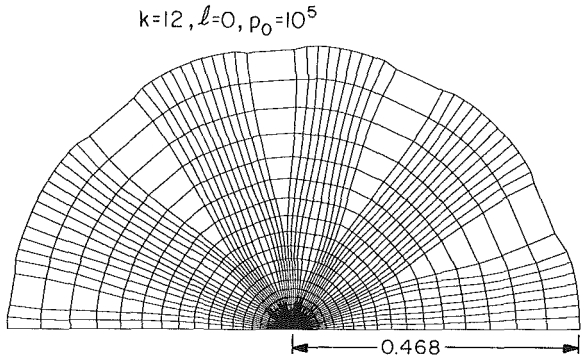


Fig. 10 Deformed mesh for a case in which ellipticity is lost

fore applies. Let C_ρ be a semicircle of radius ρ in the upper half-plane, embedded in the deformed body and centered at the origin. Let $f(\rho)$ denote the total force in the vertical ($\phi = \pi/2$) direction exerted by the material inside of C_ρ on the material outside of it. By direct integration of the traction field associated with equation (18) one finds that

$$f(\rho) \to \begin{cases} \infty, & \text{if } m \le -1 \\ 0, & \text{if } m > -1 \end{cases} \text{ as } \rho \to 0. \tag{25}$$

This result for $m \le -1$ is consistent with the fact that the tractions along the horizontal axis are tensile except at the origin, where the traction is not defined. A numerical solution for a material in which $m \le -1$ is discussed in the last section of this paper. The computation shows that even if a zero-traction boundary condition is applied to the initially flat surface, the tractions along the positive horizontal axis are still tensile.

As an example of the above solution for a physically reasonable material, consider the Blatz-Ko material ($k=2$, $l=0$). One finds from the above analysis that $n=0.838825$, $m=-0.644699$, and the necessary confining pressure is $p_0 \ge 3.62903a$.

2 Ellipticity and Stability

For the Blatz-Ko material ($k=2$, $l=0$), Knowles and Sternberg (1975) showed that the equilibrium equation can change in type from elliptic to hyperbolic depending on the local deformation. We now investigate possible loss of strong ellipticity in the creasing singularity for more general k and l.

Let $c_{\alpha\beta\gamma\delta}$ be the plane-strain elastic moduli:

$$c_{\alpha\beta\gamma\delta}(\mathbf{F}) = \frac{\partial^2 W}{\partial F_{\alpha\beta} F_{\gamma\delta}}(\mathbf{F}). \tag{26}$$

If at some point $\mathbf{x}$ the deformation gradient tensor is $\mathbf{F}$, then strong ellipticity holds at $\mathbf{x}$ if and only if

$$c_{\alpha\beta\gamma\delta}(\mathbf{F}) p_\alpha q_\beta p_\gamma q_\delta > 0 \quad \text{for all } \mathbf{p}, \mathbf{q} \tag{27}$$

where $\mathbf{p}$ and $\mathbf{q}$ are arbitrary unit vectors in the plane. Loss of strong ellipticity means that "imaginary wave speeds" are possible for infinitesimal waves in some direction superposed on the large deformation, indicating a loss of material stability in the sense of Hadamard (1903). We consider here only strong ellipticity under the assumption of plane strain, which means that only waves propagating in directions within the plane are considered.

In order to simplify the analysis, we note that for an isotropic hyperelastic material, equation (27) holds for $\mathbf{F}$ if and only if it holds for $\mathbf{FP}$, where $\mathbf{P}$ is an arbitrary proper orthogonal tensor. This is clear on physical grounds, since $\mathbf{P}$ merely represents a rigid rotation, and it may be shown rigorously using the identity $W(\mathbf{FP}) = W(\mathbf{F})$. For arbitrary $\mathbf{F}$, we set $\mathbf{P} = \mathbf{Q}^T$, where $\mathbf{Q}$ is obtained from the left polar decomposition $\mathbf{F} = \mathbf{VQ}$. Then the inequality in equation (27) is seen to hold for $\mathbf{F}$ if and only if it holds for $\mathbf{V}$. Recall that $\mathbf{V}$ is symmetric positive definite and is given by equation (6). It follows that strong ellipticity depends only on the principal stretches λ_1, λ_2.

Assume that $\mathbf{F}$ is symmetric positive definite and let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be a basis such that the components of $\mathbf{F}$ in this basis are $[\mathbf{F}]^e = \text{diag}(\lambda_1, \lambda_2)$. For the material in equation (11), if $\lambda_1 \ne \lambda_2$, the nonzero elastic moduli in this basis are as follows:

$$c_{1111} = \frac{a}{\lambda_1^2}[k(k+1)\lambda_1^{-k}\lambda_2^{-l} + l(l+1)\lambda_1^{-l}\lambda_2^{-k}],$$

$$c_{2222} = \frac{a}{\lambda_2^2}[l(l+1)\lambda_1^{-k}\lambda_2^{-l} + k(k+1)\lambda_1^{-l}\lambda_2^{-k}],$$

$$c_{1122} = c_{2211} = \frac{akl}{\lambda_1\lambda_2}[\lambda_1^{-k}\lambda_2^{-l} + \lambda_1^{-l}\lambda_2^{-k}] + a(k+l),$$

$$c_{1212} = c_{2121} = \frac{a(l-k)}{\lambda_1^2 - \lambda_2^2}[\lambda_1^{-k}\lambda_2^{-l} - \lambda_1^{-l}\lambda_2^{-k}],$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$c_{1221}=c_{2112}=\frac{a}{\lambda_1^2-\lambda_2^2}\left[\left(l\frac{\lambda_1}{\lambda_2}-k\frac{\lambda_2}{\lambda_1}\right)\lambda_1^{-k}\lambda_2^{-l}+\left(k\frac{\lambda_1}{\lambda_2}-l\frac{\lambda_2}{\lambda_1}\right)\lambda_1^{-l}\lambda_2^{-k}\right]-a(k+l). \tag{28}$$

If $\lambda_1 = \lambda_2 \equiv \lambda$, one finds

$$c_{1212}=c_{2121}=a(k-l)^2\lambda^{-k-l-2}/2,$$

$$c_{1221}=c_{2112}=a[(k-l)^2+2(k+l)]\lambda^{-k-l-2}/2-a(k+l), \tag{29}$$

and the other constants may be found from equation (28). Let

$$t=\lambda_1/\lambda_2. \tag{30}$$

For the special case $l=0$, a condition on the principal stretches for strong ellipticity may be derived by explicit evaluation of equation (27). The result is that strong ellipticity holds for $l=0$ if and only if

$$(k+1)\left[t+\frac{1}{t}\right]-\left[t^{k+1}+\left(\frac{1}{t}\right)^{k+1}\right]>0. \tag{31}$$

Evidently, for $l=0$, strong ellipticity holds provided that the ratio of the principal stretches is within a certain interval which depends on k. The lower endpoint of this interval will be called the *critical stretch ratio*, denoted $t_{\text{crit}}(k, 0)$. From the structure of equation (31) we see that strong ellipticity is lost if and only if $t<t_{\text{crit}}$ or $t>1/t_{\text{crit}}$. The dashed curve in Fig. 5 shows $t_{\text{crit}}(k, 0)$ as a function of k. Note that t_{crit} increases monotonically with increasing k and approaches 1 in the limit $k\to\infty$. A special case is the Blatz-Ko material ($k=2$, $l=0$), whose ellipticity was analyzed by Knowles and Sternberg (1975). Their results showed that $t_{\text{crit}}\approx 0.2679$ for this material in plane strain, which is identical to the result that is obtained from equation (31).

Returning to the creasing problem, equation (15) shows that in the creasing singularity the ratio of the principal stretches is independent of R and is given by

$$t_*(k,l)=\frac{\lambda_1}{\lambda_2}=\frac{\lambda_R}{\lambda_\phi}=\frac{1}{2n}, \tag{32}$$

where n is the constant found from equation (16). The solid line in Fig. 5 shows $t_*(k, 0)$ as a function of k. The conclusion is that for $l=0$, the creasing singularity undergoes loss of strong ellipticity whenever $k>9.1901$, which is the point of intersection of the curves.

For the more general case of possibly nonzero l, a condition analogous to equation (31) shows that the ellipticity of the material once again depends only on t. However, the condition is quite complicated, and the $l\neq 0$ case is handled more easily using numerical methods to search for zero wave speeds for various values of t. As in the $l=0$ case, one finds that for each k and l there is a number $t_{\text{crit}}(k, l)$ such that ellipticity is lost if and only if $t<t_{\text{crit}}$ or $t>1/t_{\text{crit}}$. After evaluating t_* using equations (15) and (32) for arbitrary k and l, it is straightforward to find the domain in the (k, l)-plane in which ellipticity is lost. This domain is shown in Fig. 6.

It should be noted that the three-dimensional analogue of equation (27) may yield a condition for strong ellipticity that is significantly different from the plane-strain condition (Knowles and Sternberg, 1975).

3 Discussion

It has been shown that any compressible elastic material in the family considered here sustains a singular deformation of the form equation (14). For physically reasonable choices of k and l, this singularity meets all requirements for physical reasonableness, including finite total strain energy. Certain combinations of the material constants k and l correspond to materials in which the requirement of compressive traction along the flat surface cannot be met regardless of p_0; hence, the singularity does not model creasing in these cases. However, such materials all have $l<0$, and their physical applicability is therefore questionable for the reasons discussed in the remarks following equation (13). For the other materials, with $l\geq 0$, equation (14) is a solution to the creasing problem for sufficiently large confining pressures (see equation (24)). Strong ellipticity in plane strain is lost for certain combinations of the constants k and l.

Full two-dimensional finite difference calculations of the creasing problem for a semicircular cylinder were carried out in order to find out what happens in the special cases mentioned previously, in which equation (14) does not apply. The computer program used was CHIMP (Silling, 1988). Only the upper half-plane was modeled, and the undeformed body was a quarter circle with radius $r_0=1$. The numerical results agree very closely with the analytic solution discussed previously when the material remains elliptic and when the confining pressure p_0 exceeds $p_{\min}(k,l)$ defined in equation (24). A typical case is shown in Fig. 7, which is for the Blatz-Ko material ($k=2$, $l=0$, $a=1$) with $p_0=4$. In this material $p_{\min}=3.62903$. For the same material with $p_0=2$, the deformed mesh is shown in Fig. 8. (The part of the initially flat surface which is not in contact with the horizontal axis is traction-free.)

An example of a material in which $m<-1$, and the first of equation (24) is therefore violated, is $k=2$, $l=-1$, $a=1$, $p_0=10$, which is shown in Fig. 9. In this case the computation predicts that $\tau_{\phi\phi}>0$ along the *positive* horizontal axis. This is true even though the initially flat surface fails to make contact with the axis, and is therefore traction-free. (See the discussion near the end of Section 1.)

An example of loss of strong ellipticity is shown in Fig. 10, which is for the material $k=12$, $l=0$, $a=1$, with $p_0=10^5$. The initial condition is the analytic solution derived above, which is an elastically unstable solution to the boundary value problem when strong ellipticity is lost. The localized strains develop spontaneously. The locations of these localizations are essentially random, since they develop from numerical noise within the mesh. This noise is amplified because of the presence of "imaginary wave speeds." The calculation does not attain an equilibrated solution, since the localized strains in this material represent failure, and the body deforms without limit.

Acknowledgment

The author gratefully acknowledges the support of the National Science Foundation through grant MSM-8658107 and through a grant of computer resources at the John von Neumann National Supercomputer Center.

References

Blatz, P. J., and Ko, W. L., 1962, "Application of Finite Elasticity Theory to the Deformation of Rubbery Materials," *Trans. Soc. Rheol.*, Vol. 6, pp. 223–251.

Carroll, M. M., 1988, "Finite Strain Solutions in Compressible Isotropic Elasticity," *J. Elasticity*, Vol. 20, pp. 65–92.

Hadamard, J., 1903, *Lecons sur la Propagation des Ondes et les Équations de l'Hydrodynamique*, Hermann, Paris, p. 241.

Singh, M., and Pipkin, A. C., 1965, "Note on Ericksen's Problem," *J. Appl. Math. Phys. (ZAMP)*, Vol. 16, pp. 706–709.

Knowles, J. K., and Sternberg, E., 1975, "On the Ellipticity of the Equations of Nonlinear Elastostatics for a Special Material," *J. Elasticity*, Vol. 5, pp. 341–361.

Silling, S. A., 1988, "Finite Difference Modeling of Phase Changes and Localization in Elasticity," *Comp. Meth. Appl. Mech. Engin.*, Vol. 70, pp. 251–273.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

An Edge Dislocation in a Three-Phase Composite Cylinder Model

H. A. Luo[1]
Graduate Student.

Y. Chen
Fellow ASME

Department of Mechanics and Materials Science,
Rutgers–The State University of New Jersey,
Piscataway, NJ 08854

An exact solution is given for the stress field due to an edge dislocation embedded in a three-phase composite cylinder. The force on the dislocation is then derived, from which a set of simple approximate formulae is also suggested. It is shown that, in comparison with the two-phase model adopted by Dundurs and Mura (1964), the three-phase model allows the dislocation to have a stable equilibrium position under much less stringent combinations of the material constants. As a result, the so-called trapping mechanism of dislocations is more likely to take place in the three-phase model. Also, the analysis and calculation show that in the three-phase model the orientation of Burgers vector has only limited influence on the stability of dislocation. This behavior is pronouncedly different from that predicted by the two-phase model.

1 Introduction

The interaction of dislocations with inhomogeneities is an important topic in studying the strengthening and hardening mechanism of many materials. This is mainly due to the fact that the mobility of dislocation can be pronouncedly influenced by various inhomogeneities in materials. In the absence of applied force and residual stresses, this interaction arises from two sources: the misfit in the atomic arrangements at the interfaces and the mismatch in the elastic properties (Dundurs, 1967). When the dislocations are not extremely close to the interface and the size of inclusion is fairly large, the elastic interaction is expected to be predominant (Dundurs, 1967).

In view of its importance, this problem has received much attention during the last several decades. Head (1953a, 1953) first derived the force on the dislocation which is near the straight interface between two dissimilar media. He found that the dislocation is either repelled or attracted by the interface, depending on the combination of material constants. This phenomenon has been observed in some other cases. Peculiarly, when Dundurs and Mura (1964) studied the interaction between an edge dislocation and a circular inclusion, they found that under certain conditions the dislocation may even have a stable equilibrium position at a distance from the interface. This result suggests a so-called trapping mechanism of the motion of dislocations near inhomogeneities. However, at the same time, Dundurs and Mura's analysis indicates that existence of a stable equilibrium position is possible only for a restricted class of materials. For example, it exists for an edge dislocation with a radial Burgers vector only when $1/\sqrt{5} < \mu_1/\mu_2 < 1$ and for an edge dislocation with a tangential Burgers vector only when $\mu_1/\mu_2 > 1$ where μ_1 and μ_2 are the shear moduli of inclusion and matrix, respectively, and for the corresponding quite limited combinations of the Poisson's ratios.

So far, most work on the dislocation was concerned with an isolated inhomogeneity. It is noted that most of the materials for engineering applications are multiphase systems. And for two-phase materials, when the inclusion phase has finite concentration, dislocation interacts not only with the nearest inclusion but also with the surrounding ones. However, to thoroughly probe the stress and strain field based on all details of microstructure would be prohibitive. Fortunately, there is a reasonably simplified model to reflect the mean effect of these interactions. In a two-dimensional case the model consists of three concentric regions: the inner circular region representing the inclusion phase, the intermediate annular region representing the matrix phase and the infinitely extended outer region representing the composite phase, which can be considered as an equivalent homogeneous transversely isotropic medium characterized by five independent effective elastic moduli of the two-phase composite. This model is the so-called three-phase cylinder model introduced by Christenson and Lo (1979). In the following analysis, we shall assume that an edge dislocation is located in the intermediate matrix phase.

By means of the three-phase cylinder model, we shall determine the stress field in the solid. In Section 2, with the aid of the analytical continuation technique of complex potentials (Muskhelishvili, 1953), the boundary value problem will be reduced to two inhomogeneous Hilbert problems. An exact solution then will be obtained by solving the Hilbert problems in an algebraic procedure. The stress field will be given in Section 3.

Based on the stress distribution in the body, the strain energy as well as the force acting on dislocation will be derived in Section 4. In Section 5, the equilibrium of dislocation within

[1]Currently, Associate Professor, Department of Engineering Mechanics, Shanghai Jiao-tung University, Shanghai, China.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 9, 1989; final revision, Mar. 1, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

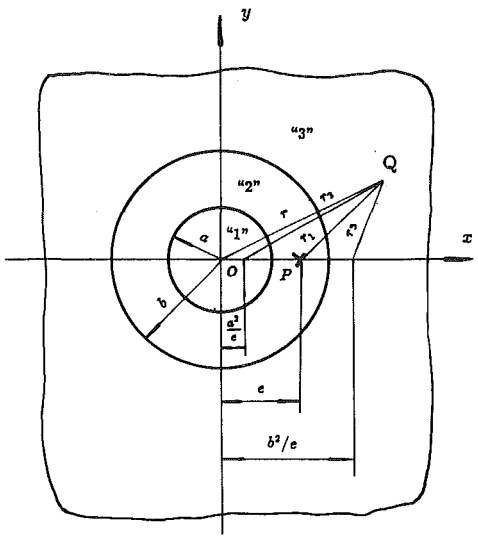


Fig. 1 Edge dislocation in the three-phase cylinder composite model

the annular intermediate region will be discussed in detail. In addition to the exact solution, an approximate expression of the force on dislocation will be suggested for engineering application. The relations between the force on dislocation and different parameters, such as elastic constant, characteristic lengths, and the volume fraction of inclusions in two-phase composite materials, will also be illustrated, respectively, for several selected material combination in that section.

2 Formulation

In what follows we will confine ourselves to the plane deformation of transversely isotropic elastic solids. For a region bounded by a circle, say, $|z| = c$, its stress and displacement fields can be expressed in terms of Muskhelishvili's complex potentials as

$$\sigma_{rr}+\sigma_{\theta\theta}=2[\Phi(z)+\overline{\Phi(z)}], \tag{1}$$

$$\sigma_{rr}+i\sigma_{r\theta}=\Phi(z)-\Phi\left(\frac{c^2}{\bar{z}}\right)+\bar{z}\left(\frac{\bar{z}}{c^2}-\frac{1}{z}\right)\overline{\Psi(z)}, \tag{2}$$

$$2\mu(u'+v')=iz\left[\kappa\Phi(z)+\Phi\left(\frac{c^2}{\bar{z}}\right)-\bar{z}\left(\frac{\bar{z}}{c^2}-\frac{1}{z}\right)\overline{\Psi(z)}\right] \tag{3}$$

in polar coordinates, where $z = x + iy$, $\bar{z} = x - iy$, $u' = \partial u_x/\theta$, $v' = \partial u_y/\partial\theta$, $\kappa = 1 + 2\mu_{12}/K_{12}$, μ is the in-plane shear modulus, namely μ_{12}, and K_{12} is the plane-strain bulk modulus. In addition, the potential $\Psi(z)$ is related to $\Phi(z)$ as

$$\Psi(z)=\frac{c^2}{z^2}\left[\Phi(z)+\bar{\Phi}\left(\frac{c^2}{z}\right)-z\Phi'(z)\right] \tag{4}$$

where $\bar{\Phi}(c^2/z) = \overline{\Phi(c^2/\bar{z})}$.

Consider the problem shown in Fig. 1. Medium 1, with elastic properties κ_1 and μ_1, occupies the inner circular region $r \le a$; medium 2, with elastic properties κ_2 and μ_2, occupies the intermediate annular region $a \le r \le b$; medium 3, with elastic properties κ_3 and μ_3, occupies the outer infinite region $r \ge b$. All interfaces are assumed to be perfectly bonded. The elastic system is free of external applied load, but with an edge dislocation $\mathbf{b} = (b_x, b_y)$ located at point $P(e, 0)$ within medium 2.

The continuity of traction and displacement on interfaces $(\sigma_{rr} + i\sigma_{r\theta})_{(i)} = (\sigma_{rr} + i\sigma_{r\theta})_{(i+1)}$ and $(u_x + iu_y)_{(i)} = (u_x + iu_y)_{i+1}$ leads to

$$\Phi^+_{(i)}(\tau)-\Phi^-_{(i)}(\tau)=\Phi^-_{(i+1)}(\tau)-\Phi^+_{(i+1)}(\tau)$$

or

$$[\Phi_{(i)}(\tau)+\Phi_{(i+1)}(\tau)]^+=[\Phi_{(i)}(\tau)+\Phi_{(i+1)}(\tau)]^- \tag{5}$$

and

$$\frac{\kappa_i}{2\mu_i}\Phi^+_{(i)}(\tau)+\frac{1}{2\mu_i}\Phi^-_{(i)}(\tau)=\frac{\kappa_{i+1}}{2\mu_{i+1}}\Phi^-_{(i+1)}(\tau)+\frac{1}{2\mu_{i+1}}\Phi^+_{(i+1)}(\tau) \tag{6}$$

for $i = 1$, $|\tau| = a$, and $i = 2$, $|\tau| = b$ where the superscripts "+" and "−" denote the value of function obtained as z approaches to the interfaces from the interior and exterior, respectively. According to the continuation theorem of the analytical function and by reference to the structure of $\Phi_{(1)}(z)$ and $\Phi_{(2)}(z)$, it yields that the function $\Xi_1(z) = \Phi_{(1)}(z) + \Phi_{(2)}(z)$ is holomorphic in the region $a^2/b < |z| < b$, except at the poles $z = e$ and $z = a^2/e$ where its principal parts are

$$\Xi_1\left(\frac{a^2}{e}\right)\sim-\frac{\gamma_{-1}}{z-\frac{a^2}{e}}+\frac{\bar{\gamma}_{-1}a^2(e^2-a^2)}{e^3}\frac{1}{\left(z-\frac{a^2}{e}\right)^2},$$

$$\Xi_1(e)\sim\frac{\gamma_{-1}}{z-e} \tag{7}$$

with $\gamma_{-1} = -i\mu_2(b_x + ib_y)/\pi(1 + \kappa_2)$ where (b_x, b_y) is the Burgers vector of the dislocation.

Therefore, we can express

$$\Xi_1(z)=-\overline{\Phi_{(1)}(0)}+\frac{\gamma_{-1}}{z}+\frac{\gamma_{-1}}{z-e}-\frac{\gamma_{-1}}{z-\frac{a^2}{e}}+\frac{\bar{\gamma}_{-1}a^2(e^2-a^2)}{e^3}\times\frac{1}{\left(z-\frac{a^2}{e}\right)^2}+\sum_{n=-\infty}^{\infty}A_nz^n \tag{8}$$

where the undetermined Laurent series is convergent in the region $a^2/b < |z| < b$. By using the Liouville theorem, it is readily shown that the sum of the first five terms in the right-hand side of equation (8) represents the summation $\Phi_{(1)}(z) + \Phi_{(2)}(z)$ when phase 2 and phase 3 are of the same elastic constants. Thus, the infinite series in this expression, the last term of (8), reflects the perturbation of $\Xi_1(z)$ due to the mismatch in the elastic properties between phase 2 and phase 3.

Eliminating $\Phi^+_{(2)}(\tau)$ and $\Phi^-_{(2)}(\tau)$ with the relation $\Phi_{(1)}(\tau) + \Phi_{(2)}(\tau) = \Xi_1(\tau)$ and using (8), we reduce (6) with $i = 1$ to the first Hilbert problem

$$\Phi^+_{(1)}(\tau)+\eta_1\Phi^-_{(1)}(\tau)=f_1(\tau) \tag{9}$$

for $|\tau| = a$, where

$$f_1(\tau)=\beta_1\Xi_1(\tau),\ \eta_1=\frac{\kappa_2\mu_1+\mu_2}{\kappa_1\mu_2+\mu_1},\ \beta_1=\frac{\mu_1(\kappa_2+1)}{\kappa_1\mu_2+\mu_1}. \tag{10}$$

Introduce a sectionally holomorphic function

$$\Omega_1(z)=\begin{cases}\Phi_{(1)}(z) & \text{when } |z|<a\\ -\eta_1\Phi_{(1)}(z) & \text{when } |z|>a.\end{cases} \tag{11}$$

Equation (9) then becomes $\Omega_1^+(\tau) - \Omega_1^-(\tau) = f_1(\tau)$ for $|\tau| = a$. Since $\lim_{z\to\infty}\Phi_{(1)}(z) = -\overline{\Phi_{(1)}(0)}$, we have

$$\Omega_1(z)=\frac{1}{2\pi i}\oint_{|\tau|=a}\frac{f_1(\tau)}{\tau-z}\,d\tau+\eta_1\overline{\Phi_{(1)}(0)}. \tag{12}$$

Substituting (12) into (11), after integration we obtain

$$\Phi_{(1)}(z)=-(\beta_2-\eta_1)\overline{\Phi_{(1)}(0)}+\frac{\gamma_{-1}\beta_1}{z-e}+\beta_1\sum_{n=0}^{\infty}A_nz^n,$$

(for $|z| < a$)

and

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\Phi_{(1)}(z)=\frac{\beta_1}{\eta_1}\left[\frac{\gamma_{-1}}{z}-\frac{\gamma_{-1}}{z-\frac{a^2}{e}}+\frac{\overline{\gamma}_{-1}a^2(e^2-a^2)}{e^3}\frac{1}{\left(z-\frac{a^2}{e}\right)^2}+\sum_{n=1}^{\infty}A_n z^{-n}\right]-\overline{\Phi_{(1)}(0)}. \quad (13)$$

(for $|z|>a$)

with

$$\overline{\Phi_{(1)}(0)}=\frac{\beta_1\overline{A}_0-\beta_1(\beta_1-\eta_1)A_0}{1-(\beta_1-\eta_1)^2}-\frac{\overline{\gamma}_{-1}\beta_1-\gamma_{-1}\beta_1(\beta_1-\eta_1)}{e[1-(\beta_1-\eta_1)^2]}. \quad (14)$$

Subtracting equation (13) from $\Xi_1(z)$, we obtain

$$\Phi_{(2)}(z)=-(1-\beta_1+\eta_1)\overline{\Phi_{(1)}(0)}+(1-\beta_1)\left(\frac{\gamma_{-1}}{z-e}+\sum_{n=0}^{\infty}A_n z^n\right)+\frac{\gamma_{-1}}{z}-\frac{\gamma_{-1}}{z-\frac{a^2}{e}}+\frac{\overline{\gamma}_{-1}a^2(e^2-a^2)}{e^3}\frac{1}{\left(z-\frac{a^2}{e}\right)^2}+\sum_{n=1}^{\infty}A_{-n}z^{-n},$$

(for $a^2/b<|z|<a$)

and

$$\Phi_{(2)}(z)=\frac{\gamma_{-1}}{z-e}+\sum_{n=0}^{\infty}A_n z^n+\left(1-\frac{\beta_1}{\eta_1}\right)\left[\frac{\gamma_{-1}}{z}-\frac{\gamma_{-1}}{z-\frac{a^2}{e}}+\frac{\overline{\gamma}_{-1}a^2(e^2-a^2)}{e^3}\frac{1}{\left(z-\frac{a^2}{e}\right)^2}+\sum_{n=1}^{\infty}A_{-n}z^{-n}\right]. \quad (15)$$

(for $a<|z|<b$).

Substitution of (15) into (4) with $c=a$ gives

$$\Phi_{(2)}(z)=\frac{a^2}{z^2}\left\{\frac{1-\eta_1}{1+\beta_1-\eta_1}(A_0+\overline{A}_0)+\frac{\gamma_{-1}\beta_1-\overline{\gamma}_{-1}+\overline{\gamma}_{-1}\eta_1}{e(1+\beta_2-\eta_1)}+\frac{\gamma_{-1}(e^2-a^2)+\overline{\gamma}_{-1}e^2}{ea^2}+\frac{e^2(\overline{\gamma}_{-1}+2\gamma_{-1})}{a^2}\frac{1}{z-e}+\frac{\gamma_{-1}e^3}{a^2}\frac{1}{(z-e)^2}-\left[2\left(1-\frac{\beta_1}{\eta_1}\right)\gamma_{-1}+(1-\beta_1)\frac{\overline{\gamma}_{-1}a^2}{e^2}\right]\frac{1}{z-\frac{a^2}{e}}+\left(1-\frac{\beta_1}{\eta_1}\right)\frac{3\overline{\gamma}_{-1}a^2(e^2-a^2)-\gamma_{-1}a^2e^2}{e^3}\frac{1}{\left(z-\frac{a^2}{e}\right)^2}+\left(1-\frac{\beta_1}{\eta_1}\right)\frac{2\overline{\gamma}_{-1}a^4(e^2-a^2)}{e^4}\frac{1}{\left(z-\frac{a^2}{e}\right)^3}+\left(1-\frac{\beta_1}{\eta_1}\right)\frac{2\gamma_{-1}}{z}+\frac{\overline{\gamma}_{-1}}{a^2}z-\sum_{n=1}^{\infty}A_n(n-1)z^n+\sum_{n=1}^{\infty}\overline{A}_{-n}\left(\frac{z}{a^2}\right)^n+\left(1-\frac{\beta_1}{\eta_1}\right)\sum_{n=1}^{\infty}A_{-n}(n+1)z^{-n}+(1-\beta_1)\sum_{n=1}^{\infty}\overline{A}_n\left(\frac{a^2}{z}\right)^n\right\}. \quad (16)$$

(for $a<|z|<b$)

In the same manner, starting from the continuity conditions (5) and (6) with $i=2$, we arrive at the second Hilbert problem

$$\Phi^{+}_{(3)}(\tau)+\eta_2\Phi^{-}_{(3)}(\tau)=f_2(\tau) \quad (17)$$

for $|\tau|=b$ where

$$f_2(\tau)=\beta_2\Xi_2(\tau),\ \eta_2=\frac{\kappa_3\mu_2+\mu_3}{\kappa_2\mu_3+\mu_2},\ \beta_2=\frac{\mu_3(\kappa_2+1)}{\kappa_2\mu_3+\mu_2},$$

$$\Xi_2(\tau)=\frac{\gamma_{-1}}{\tau-e}-\frac{\gamma_{-1}}{\tau-\frac{b^2}{e}}+\frac{\overline{\gamma}_{-1}b^2(e^2-b^2)}{e^3}\frac{1}{\left(\tau-\frac{b^2}{e}\right)^2}+\frac{\overline{\gamma}_{-1}}{e}+\frac{\gamma_{-1}}{\tau}+\sum_{n=-\infty}^{\infty}B_n\tau^n. \quad (18)$$

Here, B_n's are the coefficients of a Laurent series, which is convergent in the region $a<|z|<b^2/a$. The solution of this problem is

$$\Phi_{(3)}(z)=\beta_2\left\{-\frac{\gamma_{-1}}{z-\frac{b^2}{e}}+\frac{\overline{\gamma}_{-1}b^2(e^2-b^2)}{e^3}\frac{1}{\left(z-\frac{b^2}{e}\right)^2}+\frac{\overline{\gamma}_{-1}}{e}+\sum_{n=0}^{\infty}B_n z^n\right\}+\frac{\gamma_{-1}}{z}$$

(for $|z|<b$)

$$\Phi_{(3)}(z)=\frac{\beta_2}{\eta_2}\left(\frac{\gamma_{-1}}{z-e}+\sum_{n=1}^{\infty}B_{-n}z^{-n}\right)-\frac{1-\beta_2}{\eta_2}\frac{\gamma_{-1}}{z}, \quad (19)$$

(for $|z|>b$).

Subtracting $\Phi_{(3)}(z)$ from $\Xi_2(z)=\Phi_{(2)}(z)+\Phi_{(3)}(z)$, we get

$$\Phi_{(2)}(z)=\frac{\gamma_{-1}}{z-e}+\sum_{n=1}^{\infty}B_{-n}z^{-n}+(1-\beta_2)\left[\frac{\overline{\gamma}_{-1}}{e}-\frac{\gamma_{-1}}{z-\frac{b^2}{e}}+\frac{\overline{\gamma}_{-1}b^2(e^2-b^2)}{e^3}\frac{1}{\left(z-\frac{b^2}{e}\right)^2}+\sum_{n=0}^{\infty}B_n z^n\right]$$

(for $a<|z|<b$)

$$\Phi_{(2)}(z)=\frac{\overline{\gamma}_{-1}}{e}-\frac{\gamma_{-1}}{z-\frac{b^2}{e}}+\frac{\overline{\gamma}_{-1}b^2(e^2-b^2)}{e^3}\frac{1}{\left(z-\frac{b^2}{e}\right)^2}+\sum_{n=0}^{\infty}B_n z^n+\left(1-\frac{\beta_2}{\eta_2}\right)\left(\frac{\gamma_{-1}}{z-e}+\sum_{n=1}^{\infty}B_{-n}z^{-n}\right)+\frac{1-\beta_2+\eta_2}{\eta_2}\left(\frac{\gamma_{-1}}{z}\right). \quad (20)$$

(for $b<|z|<b^2/a$)

And also,

$$\Psi_{(2)}(z)=\frac{b^2}{z^2}\left\{\beta_2\left(\frac{1}{\eta_2}-1\right)\frac{\overline{\gamma}_{-1}}{e}+\frac{(\gamma_{-1}+\overline{\gamma}_{-1})^e}{b^2}+\overline{B}_0+(1+\beta_2)B_0+\frac{e^2(\overline{\gamma}_{-1}+2\gamma_{-1})}{b^2}\frac{1}{z-e}+\frac{e^3\gamma_{-1}}{b^2}\frac{1}{(z-e)^2}-\left[2(1-\beta_2)\gamma_{-1}\right.\right.$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+1\left(1-\frac{\beta_2}{\eta_2}\right)\frac{\overline{\gamma}_{-1}b^2}{e^2}\Bigg]\frac{1}{z-\frac{b^2}{e}}$$

$$+(1-\beta_2)\frac{3\overline{\gamma}_{-1}b^2(e^2-b^2)-\gamma_{-1}b^2e^2}{e^3}\frac{1}{\left(z-\frac{b^2}{e}\right)^2}$$

$$+(1-\beta_2)\frac{2\overline{\gamma}_{-1}b^4(e^2-b^2)}{e^4}\frac{1}{\left(z-\frac{b^2}{e}\right)^3}$$

$$-(1-\beta_2)\sum_{n=1}^{\infty}B_n(n-1)z^n$$

$$+\left(1+\frac{\beta_2}{\eta_2}\right)\sum_{n=1}^{\infty}\overline{B}_{-n}\left(\frac{z}{b^2}\right)^n$$

$$+\sum_{n=1}^{\infty}B_{-n}(n+1)z^{-n}+\sum_{n=1}^{\infty}\overline{B}_n\left(\frac{b^2}{z}\right)^n$$

$$+\frac{1-\beta_2+\eta_2}{\eta_2}\left(\frac{\overline{\gamma}_{-1}}{b^2}\right)z\Bigg\} \tag{21}$$

(for $a<|z|<b$).

In order to satisfy the continuity conditions of displacement and traction on the interface $|z| = a$ and $|z| = b$ simultaneously, the function $\Phi_{(2)}(z)$ expressed by equations $(15)_2$ and $(20)_1$ and $\Psi_{(2)}(z)$ by (16) and (21) must be compatible, respectively (England, 1971). Thus, we have two compatibility identities. Expanding both sides of each identity into power series of z/a and comparing the coefficients of their like powers, we obtain a set of simultaneous equations with respect to unknowns $A_n^* = A_n a^n$, $B_n^* = B_n a^n$ $(n = 0, \pm 1, \pm 2, \ldots)$. If we further define

$$A_n^*=\frac{\mu_2}{\pi(1+\kappa_2)_a}(a_nb_y+ia_n'b_x),$$

$$B_n^*=\frac{\mu_2}{\pi(1+\kappa_2)_a}(b_nb_y+ib_n'b_x), \tag{22}$$

we get two sets of real simultaneous equations:

$$\frac{1}{D}a_n-b_n=\frac{1}{\beta}\delta_{n0}+\beta^{n+1}\delta^{2n+2}+(n+1)(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}$$

(for $n\geq 0$)

$$a_{-n}+\frac{1}{A}b_{-n}=\delta_{n1}-\beta^{1-n}+(n-1)(1-\beta^{-2})\beta^{1-n}$$

(for $n\geq 1$)

$$-[(n-1)-(N-1)\delta_{n0}]\delta^2a_n+[1+(N-1)\delta_{n0}]\delta^2a_{-n}+D(n-1)b_n$$

$$-[C\delta^{2n}+(1-C)\delta_{n0}]b_{-n}=\left[(D-C)\frac{1}{\beta}+\frac{2N}{\beta}\delta^2\right]\delta_{n0}$$

$$+(C+2D\beta^2\delta^2)\times\beta^{n-1}\delta^{2n}$$

$$+3D(n+1)(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}-D(n+1)\beta^{n+1}\delta^{2n+2}$$

$$-2(-1)^nD(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}C_{-3}^n+\frac{D(1-C)}{1-D}\delta^2\delta_{n1}$$

(for $n\geq 0$)

$$-A(n+1)\delta^2a_{-n}-B\delta^2a_n+(n+1)b_n+\delta^{-2n}b_n$$

$$=-A(n+1)\beta^{1-n}\delta^2-B\beta^{-(n+1)}\delta^3+3A(n-1)(\beta^2-1)\delta^2/\beta^{-(n+1)}$$

$$+2A(-1)^{n+1}(\beta^2-1)\delta^2\beta^{-(n+1)}C_{-3}^{n-3}(1-\delta_{n1})(1-\delta_{n2})$$

$$+2A\delta^2\delta_{n1} \tag{23}$$

(for $n\geq 1$)

and

$$\frac{1}{D}a_n'-b_n'=\frac{1}{\beta}\delta_{n0}-\beta^{n+1}\delta^{2n+2}+(n+1)(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}$$

(for $n\geq 0$)

$$-a_{-n}'+\frac{1}{A}b_{-n}'=-\delta_{n1}+\beta^{1-n}+(n-1)(1-\beta^{-2})\beta^{1-n}$$

(for $n\geq 1$)

$$-[(n-1)-(N-1)\delta_{n0}]\delta^2a_n'-[1+(N-1)\delta_{n0}]\delta^2a_{-n}'+D(n-1)b_n'$$

$$+[C\delta^{2n}+(1-C)\delta_{n0}]b_{-n}'=(D-C)\frac{1}{\beta}\delta_{n0}$$

$$+(C-2D\beta^2\delta^2)\beta^{n-1}\delta^{2n}+3D(n+1)(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}$$

$$+D(n+1)\beta^{n+1}\delta^{2n+2}-2(-1)^nD(\beta^2\delta^2-1)\beta^{n-1}\delta^{2n}C_{-3}^n$$

$$+\frac{D(1-C)}{1-D}\delta^2\delta_{n1}$$

(for $n\geq 0$)

$$-A(n+1)\delta^2a_{-n}'+B\delta^2a_n'+(n+1)b_{-n}'-\delta^{-2n}b_n'$$

$$=A(n+1)\beta^{1-n}\delta^2-B\beta^{-(n+1)}\delta^2+3A(n-1)(\beta^2-1)\delta^2\beta^{-(n+1)}$$

$$+2A(-1)^{n+1}(\beta^2-1)\delta^2\beta^{-(n+1)}C_{-3}^{n-3}(1-\delta_{n1})(1-\delta_{n2})$$

$$-2A\delta^2\delta_{n1} \tag{24}$$

(for $n\geq 1$),

where δ_{nl} is the Kronecker delta, and

$$\beta=\frac{e}{a},\quad \delta=\frac{a}{b},\quad m=\frac{\mu_1}{\mu_2},\quad l=\frac{\mu_3}{\mu_2},$$

$$A=1-\frac{\beta_1}{\eta_1}=\frac{1-m}{1+m\kappa_2},\quad B=1-\beta_1=\frac{\kappa_1-m\kappa_2}{\kappa_1+m},$$

$$C=1-\frac{\beta_2}{\eta_2}=\frac{\kappa_3-l\kappa_2}{\kappa_3+l},\quad D=1-\beta_2=\frac{1-l}{1+l\kappa_2},$$

$$N=\frac{1-\eta_1}{1+\beta_2-\eta_1}=\frac{\kappa_1-1-m(\kappa_2-1)}{\kappa_1-1+2m}. \tag{25}$$

3 Stress Distribution

In this problem, of particular interest is the stress field. In region 2, i.e., in the intermediate medium, we have

$$\sigma_{xx}^{(2)}=\mathrm{Re}[2\Phi_{(2)}(z)-\overline{z}\Phi_{(2)}'(z)-\Psi_{(2)}(z)]$$

$$=\frac{\mu_2}{\pi(1+\kappa_2)}[h_{xx1}^{(2)}b_x+h_{xx2}^{(2)}b_y],$$

$$\sigma_{yy}^{(2)}=\mathrm{Re}[2\Phi_{(2)}(z)+\overline{z}\Phi_{(2)}'(z)+\Psi_{(2)}(z)]$$

$$=\frac{\mu_2}{\pi(1+\kappa_2)}[h_{yy1}^{(2)}b_x+h_{yy2}^{(2)}b_y],$$

$$\sigma_{xy}^{(2)}=\mathrm{Im}[\overline{z}\Phi_{(2)}'(z)+\Psi_{(2)}(z)]$$

$$=\frac{\mu_2}{\pi(1+\kappa_2)}[h_{xy1}^{(2)}b_x+h_{xy2}^{(2)}b_y], \tag{26}$$

where the influence functions $h_{xx1}^{(2)}$, $h_{xx2}^{(2)}$, $h_{yy1}^{(2)}$, $h_{yy2}^{(2)}$, $h_{xy1}^{(2)}$, and $h_{xy2}^{(2)}$, are listed in the Appendix.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

It is noticed that if the materials in the intermediate phase and the outer phase are of the same elastic properties, i.e., $\kappa_2 = \kappa_3$, $\mu_2 = \mu_3$, then in equation (25) we have $C = D = 0$. Thus, from $(23)_1$, $(23)_3$, $(24)_1$, and $(24)_3$ we obtain $a_n = a'_n = 0$ for $-\infty < n < \infty$. Therefore, the stress field is reduced to that derived by Dundurs and Mura (1964) for an infinite body embedded by a circular inclusion with an edge dislocation in outer phase (Erdogan et al., 1974).

In the same manner, we can determine the stress field in phase 1 and 3. Here we omit the details.

4 Force on Dislocation

Following Dundurs and Mura (1964), we can calculate the strain energy of the three-phase solid by evaluating the work required to introduce a dislocation into the given location. Thus, for a dislocation with component b_x only, the strain energy per unit thickness of the body is

$$W_1 = \frac{1}{2} b_x \int_{e+r_0}^{b} \sigma_{xy}^{(2)}(x,0)dx + \frac{1}{2} b_x \int_{b}^{R} \sigma_{xy}^{(3)}(x,0)dx \quad (27)$$

where r_0 is the radius of the core of the dislocation, and R is the outer radius of the body. It is relevant to assume $r_0 \to 0$ and $R \to \infty$. By substitution we obtain

$$\begin{aligned}
W_1 = \frac{\mu_2 b_x^2}{2\pi(1+\kappa_2)}\Bigg\{ & 2\log\frac{b-e}{r_0} + (B+A)\log\frac{\beta^2-1}{\beta^2} - (3A-B)\frac{1}{\beta^2} \\
& - (A+B)\log\left(1-\frac{\delta}{\beta}\right) + 2A\frac{\beta^2-1}{\beta^2}\frac{\delta}{\beta-\delta} \\
& - A\frac{(\beta^2-1)^2}{\beta^2}\frac{\delta^2}{(\beta-\delta)^2} - (B-A)\frac{\delta}{\beta} + A\delta^2 \\
& - \sum_{n=2}^{\infty}\left(a'_n + \frac{1}{n-1}a'_{-n}\right)(\delta^{1-n} - \beta^{n-1}) \\
& + \sum_{n=1}^{\infty} a'_n \frac{n}{n+1}[\delta^{-(n+1)} - \beta^{n+1}] \\
& + A\sum_{n=2}^{\infty} a'_{-n}\frac{n}{n-1}(\delta^{n-1} - \beta^{1-n}) \\
& - \sum_{n=1}^{\infty}\left(Aa'_{-n} - \frac{B}{n+1}a'_n\right)[\delta^{n+1} - \beta^{-(n+1)}] \\
& + (1+A)a'_{-1}\log(\delta\beta) + (2-C-D)\log\frac{R}{b-e} \\
& + \frac{D(C-D)}{1-D}\log\frac{R}{b} \\
& + \frac{D(1-C)}{1-D} - (D-C) - \frac{1-D}{\delta}b'_0 \\
& - (1-C)b'_{-1}\log\frac{R}{b} - (1-C)\sum_{n=2}^{\infty} b'_{-n}\frac{n}{n-1}\delta^{n-1} \\
& - (2-D)\sum_{n=1}^{\infty} b'_n \frac{1}{n+1}\delta^{-(n+1)} \\
& + (1-C)\sum_{n=1}^{\infty} b'_{-n}\delta^{n-1}\Bigg\}. \quad (28)
\end{aligned}$$

Similarly, we get the strain energy per unit thickness due to dislocation b_y:

$$\begin{aligned}
W_2 = \frac{\mu_2 b_y^2}{2\pi(1+\kappa_2)}\Bigg\{ & 2\log\frac{b-e}{e_0} + (A+B)\log\frac{\beta^2-1}{\beta^2} \\
& + \frac{1}{\beta^2}(A+B-2N) - (A+B)\log\left(1-\frac{\delta}{\beta}\right) + \frac{\delta}{\beta}(A-B+2N) \\
& - A\delta^2 + 2A\frac{\delta(\beta^2-1)}{\beta-\delta} \\
& - A\frac{(\beta^2-1)^2}{\beta^2}\frac{\delta^2}{(\beta-\delta)^2} - 2A\delta\beta - 2Na_0\left(\delta - \frac{1}{\beta}\right) \\
& + \sum_{n=0}^{\infty} a_n\left(\frac{n+2}{n+1}\right)[\delta^{-(n+1)} - \beta^{n+1}] \\
& - \sum_{n=2}^{\infty}\left(a_n - \frac{1}{n-1}a_{-n}\right)(\delta^{1-n} - \beta^{n-1}) \\
& + A\sum_{n=2}^{\infty} a_{-n}\left(\frac{n-2}{n-1}\right)(\delta^{n-1} - \beta^{1-n}) \\
& - \sum_{n=1}^{\infty}\left[Aa_{-n} + \frac{B}{n+1}a_n\right][\delta^{n+1} - \beta^{-(n+1)}] \\
& - (A+1)a_{-1}\log(\delta\beta) + (2-C-D)\log\frac{R}{b-e} \\
& + \frac{D(C-D)}{1-D}\log\frac{R}{b} - \frac{D(1-C)}{1-D} + (D-C) + \frac{1-D}{\delta}b_0 \\
& - (1-C)\sum_{n=2}^{\infty} b_{-n}\left(\frac{n-2}{n-1}\right)\delta^{n+1} \\
& + (1-C)b_{-1}\log\frac{R}{b} + (1-D)\sum_{n=1}^{\infty} b_n\frac{1}{n+1}\delta^{-(n+1)} \\
& + (1-C)\sum_{n=1}^{\infty} b_{-n}\delta^{n-1}\Bigg\}. \quad (29)
\end{aligned}$$

Assume that both of the foregoing two dislocations can move only in a radial direction. Then, the force on dislocation b_x, defined by the decrease of strain energy in the solid for it to glide a unit distance (Dundurs and Mura, 1964; Dundurs 1967), is

$$F_1 = -\frac{\partial W_1}{\partial e} = -\frac{1}{a}\frac{\partial W_1}{\partial \beta}. \quad (30)$$

Using equation (28), it follows that

$$\begin{aligned}
F_1 = -\frac{\mu_2 b_x^2}{\pi(1+\kappa_2)a}\Bigg\{ & \frac{A+B}{\beta(\beta^2-1)} + \frac{3A-B}{\beta^3} - (A+B)\frac{\delta}{2\beta(\beta-\delta)} \\
& - A\frac{\delta(\beta^3-3\beta+2\delta)}{\beta^3(\beta-\delta)^2} + A\frac{\delta^2(\beta^2-1)(\beta^2\delta-2\beta+\delta)}{\beta^3(\beta-\delta)^3} \\
& + (B-A)\frac{\delta}{2\beta^2} - \frac{1}{2}\sum_{n=1}^{\infty} a'_n n\beta^n + \frac{1}{2}\sum_{n=1}^{\infty}[a'_n(n-1) + a'_{-n}]\beta^{n-2} \\
& + \frac{A}{2}\sum_{n=1}^{\infty} a'_{-n}n\beta^{-n} \\
& - \frac{1}{2}\sum_{n=1}^{\infty}[A(n+1)a'_{-n} - Ba'_n]\beta^{-(n+2)} - \frac{1}{2}(C+D)\frac{\delta}{1-\beta\delta} \\
& + \frac{1}{2}\sum_{n=1}^{\infty}\frac{\partial a'_n}{\partial\beta}\frac{n}{n+1}[\delta^{-(n+1)} - \beta^{n+1}]
\end{aligned}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-\frac{1}{2}\sum_{n=2}^{\infty}\left(\frac{\partial a'_n}{\partial\beta}+\frac{1}{n-1}\frac{\partial a'_{-n}}{\partial\beta}\right)(\delta^{1-n}-\beta^{n-1})$$

$$+\frac{A}{2}\sum_{n=1}^{\infty}\frac{\partial a'_{-n}}{\partial\beta}\frac{n}{n-1}(\delta^{n-1}-\beta^{1-n})$$

$$-\frac{1}{2}\sum_{n=1}^{\infty}\left(A\frac{\partial a'_{-n}}{\partial\beta}-\frac{B}{n+1}\frac{\partial a'_n}{\partial\beta}\right)[\delta^{n+1}-\beta^{-(n+1)}]$$

$$-\frac{1-D}{2\delta}\frac{\partial b'_0}{\partial\beta}-\frac{1-C}{2}\sum_{n=2}^{\infty}\frac{\partial b'_{-n}}{\partial\beta}\frac{n}{n-1}\delta^{n-2}$$

$$-\frac{1-D}{2}\sum_{n=1}^{\infty}\frac{\partial b'_n}{\partial\beta}\frac{1}{n+1}\delta^{-(n+1)}+\frac{1-C}{2}\sum_{n=1}^{\infty}\frac{\partial b'_{-n}}{\partial\beta}\delta^{n-1}\Bigg\} \quad (31)$$

where the quantities $\partial a'_n/\partial\beta$ and $\partial b'_n/\partial\beta$ are calculated from equation (24). In fact, taking derivative with respect to β on both sides of equation (24), we get a system of equations to determine these quantities, with the coefficient matrix unchanged. From this system of equations, we can verify that

$$\frac{\partial a'_{-1}}{\partial\beta}=0,\quad \frac{\partial b'_{-1}}{\partial\beta}=0. \quad (32)$$

Actually, these two relations come from the self-equilibrium condition of the body.

Similarly, the force on dislocation b_y, defined by the decrease of strain energy in the solid for it to climb a unit distance in radial direction is

$$F_2=-\frac{\mu_2 b_y^2}{\pi(1+\kappa_2)a}\Bigg\{\frac{A+B}{\beta(\beta^2-1)}-\frac{A+B-2N}{\beta^3}-(A+B)\frac{\delta}{2\beta(\beta-\delta)}$$

$$-\frac{\delta(A-B+2N)}{2\beta^2}+A\frac{\delta(\beta^2-2\beta\delta+1)}{(\beta-\delta)^2}$$

$$+A\frac{\delta^2(\beta^2-1)(\beta^2\delta-2\beta+\delta)}{\beta^3(\beta-\delta)^3}$$

$$-A\delta-\frac{N}{\beta^2}a_0-\frac{1}{2}\sum_{n=1}^{\infty}a_n(n+2)\beta^n$$

$$+\frac{1}{2}\sum_{n=1}^{\infty}[a_n(n-1)-a_{-n}]\beta^{n-2}$$

$$+\frac{A}{2}\sum_{n=1}^{\infty}a_{-n}(n-2)\beta^{-n}-\frac{1}{2}\sum_{n=1}^{\infty}[Ba_n+A(n+1)a_{-n}]\beta^{-(n+2)}$$

$$-N\left(\delta-\frac{1}{\beta}\right)\frac{\partial a_0}{\partial\beta}+\frac{1}{2}\sum_{n=1}^{\infty}\frac{\partial a_n}{\partial\beta}\left(\frac{n+2}{n+1}\right)[\delta^{-(n+1)}-\beta^{n+1}]$$

$$-\frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{\partial a_n}{\partial\beta}-\frac{1}{n-1}\frac{\partial a_{-n}}{\partial\beta}\right)(\delta^{1-n}-\beta^{n-1})$$

$$+\frac{A}{2}\sum_{n=1}^{\infty}\frac{\partial a_{-n}}{\partial\beta}\left(\frac{n-2}{n-1}\right)(\delta^{n-1}-\beta^{1-n})$$

$$-\frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{B}{n+1}\frac{\partial a_n}{\partial\beta}+A\frac{\partial a_{-n}}{\partial\beta}\right)[\delta^{n+1}-\beta^{-(n+1)}]$$

$$-\frac{1}{2}(C+D)\frac{\delta}{1-\beta\delta}$$

$$+\frac{1-D}{2\delta}\frac{\partial b_0}{\partial\beta}-\frac{1-C}{2}\sum_{n=1}^{\infty}\frac{\partial b_{-n}}{\partial\beta}\left(\frac{n-2}{n-1}\right)\delta^{n+1}$$

$$+\frac{1-C}{2}\sum_{n=1}^{\infty}\frac{\partial b_{-n}}{\partial\beta}\delta^{n-1}+\frac{1-D}{2}\sum_{n=1}^{\infty}\frac{\partial b_n}{\partial\beta}\frac{1}{n+1}\delta^{-(n+1)}\Bigg\}. \quad (33)$$

5 Discussion

From equations (31) and (33) it is observed that in the three-phase cylinder model, the force on dislocation is a rather complicated function of the location of dislocation β, characteristic length δ, and the material properties of the solid. Therefore, in this case it seems to be a quite difficult task to determine the equilibrium position of a dislocation where the force F_1 or F_2 vanishes. Fortunately, we can still easily judge whether a dislocation is attracted or repelled by the interfaces $r = a$ and $r = b$ when it is close to them. Indeed, for dislocations b_x and b_y which are close to interface $r = a$, namely, $\beta \to 1$, the first

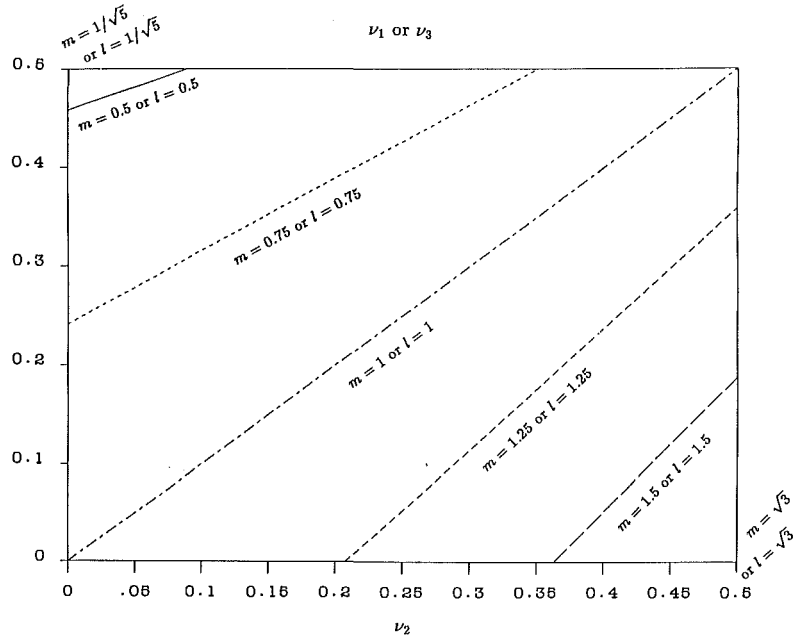


Fig. 2 Loci of $A + B = 0$ and $C + D = 0$ for a three-phase cylinder consisting of three respective isotropic media

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

terms in the right-hand side of (31) and (33) are dominant. For dislocations b_x and b_y which are close to the interface $r = b$, namely, $\beta\delta \to 1$, the term including the factor $\delta(C + D)/2(1 - \beta\delta)$ is dominant. Thus, for both dislocations b_x and b_y we have the following sets of criterion:

$$\begin{cases} A+B<0 & \text{repelled by phase 1} \\ A+B>0 & \text{attracted by phase 1} \end{cases} \tag{34}$$

for $\beta \to 1$ (or $e \to a$), and

$$\begin{cases} C+C<0 & \text{repelled by phase 3} \\ C+D>0 & \text{attracted by phase 3} \end{cases} \tag{35}$$

for $\beta \to 1/\delta$ (or $e \to b$).

Hence, for the materials which satisfy both the conditions $A + B < 0$ and $C + D < 0$, dislocations b_x and b_y will have at least a stable equilibrium position within the interval $1 \le \beta \le 1/\delta$ (or $a \le e \le b$); for the materials which satisfy both the conditions $A + B > 0$ and $C + D > 0$, dislocations b_x and b_y will have at least an unstable equilibrium position within the same interval. To be more visual, Fig. 2 shows the loci of $A + B = 0$ and $C + D = 0$ for a three-phase solid in which three respective phases are all isotropic but dissimilar. In this figure the regions $A + B < 0$ and $C + D < 0$ are located at the upper left sides of the loci $A + B = 0$ and $C + D = 0$, respectively, while the regions $A + B > 0$ and $C + D > 0$ are located at the lower right sides of the corresponding loci. Notice that in case the three phases are isotropic, the Poisson's ratios comply with $0 < \nu_i < 0.5$ where $i = 1, 2, 3$. Thus, from the above discussion it is clear that, if the shear rigidities satisfy $\mu_1/\mu_2 > \sqrt{3}$ and $\mu_3/\mu_2 > \sqrt{3}$, there must be at least a stable equilibrium position of dislocation in phase 2. If the shear rigidities satisfy $\mu_1/\mu_2 < 1/\sqrt{5}$ and $\mu_3/\mu_2 < 1/\sqrt{5}$, there must be at least an unstable equilibrium position of dislocation in that region. If either $1/\sqrt{5} < \mu_1/\mu_2 < \sqrt{3}$ or $1/\sqrt{5} < \mu_3/\mu_2 < \sqrt{3}$, whether there is an equilibrium position of dislocation, stable or unstable, will depend on the combination of the Poisson's ratios in these three phases. For instance, when all the three phases are of the same Poisson's ratios, dislocations will have a stable equilibrium position if the inclusion phase and the composition phase are harder than the matrix phase, namely, $m > 1$ and $l > 1$. However, it is not always true in the other cases. For example, in the three-phase solid with $\nu_1 = \nu_3 = 0.2$, $\nu_2 = 0.4$ even for $m = l = 1.25$, a case of the inclusion phase and the composite phase being harder than matrix phase, we have $A + B > 0$ and $C + D > 0$; the computation shows that dislocations have no stable equilibrium position. So, in the range either $1/\sqrt{5} < \mu_1/\mu_2 < \sqrt{3}$ or $1/\sqrt{5} < \mu_3/\mu_2 < \sqrt{3}$, Poisson's ratios might play some intricate role regarding the stability of dislocations.

In comparison with Dundurs and Mura's two-phase model, we can conclude that the present three-phase model allows dislocation to have a stable equilibrium position under much less stringent conditions. For example, according to Dundurs and Mura's analysis (1964), for dislocation b_x a stable equilibrium position exists only when $A + B < 0$ and $A > 0$ ($m = \mu_1/\mu_2 < 1$). That is, if $\mu_1/\mu_2 > 1$, b_x has no equilibrium position at all. However, in the three-phase model, so long as $\min(\mu_1/\mu_2, \mu_3/\mu_2) > 1/\sqrt{5}$, there exists a stable equilibrium position for some combinations of the Poisson's ratios. In particular, when $\min(\mu_1/\mu_2, \mu_3/\mu_2) > \sqrt{3}$, the stable equilibrium position exists for any combination of the Poisson's ratios. Here, we also notice that when the outer phase of the three-phase model represents the effective medium of the composite, the last condition is readily satisfied by the composites filled with hard inclusions.

Practically, if the thickness of the intermediate phase 2 is not too large, namely, the volume fraction of inclusions is not too small, we can use the following first-order approximate to roughly estimate the force on dislocation:

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = -\frac{\mu_2}{\pi(1+\kappa_2)a}\left[\frac{A+B}{\beta(\beta^2-1)} - \frac{1}{2}(C+D)\frac{\delta}{1-\beta\delta}\right]\begin{pmatrix} b_x^2 \\ b_y^2 \end{pmatrix}. \tag{36}$$

Here, only the singular terms in the right-hand sides of equations (31) and (33) are kept.

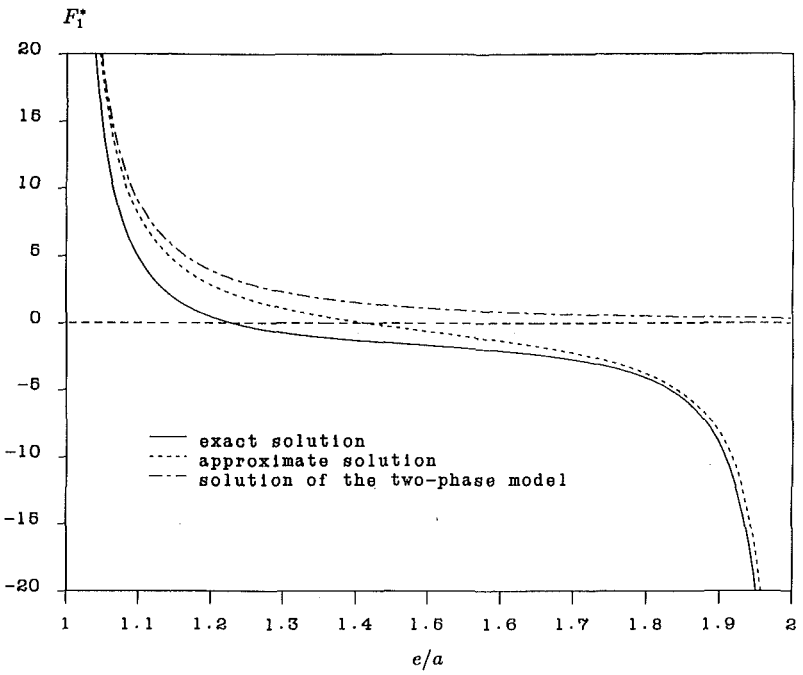


Fig. 3 Dimensionless force on the dislocation b_x in the three-phase cylinder with $\mu_1 = 20\mu_2$, $\mu_3 = 8\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$ and $a/b = 0.5$ characterized by $A + B < 0$ and $C + D < 0$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

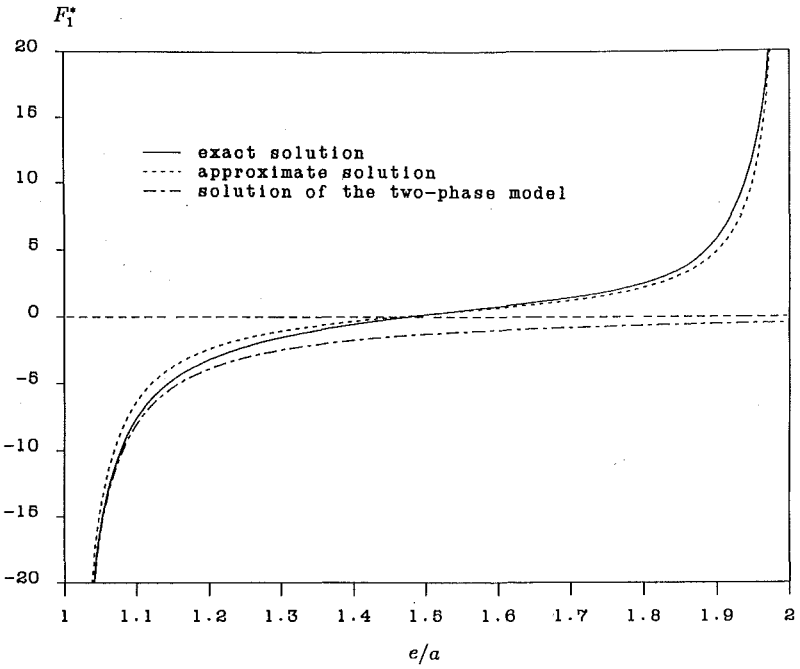


Fig. 4 Dimensionless force on the dislocation b_x in the three-phase cylinder with $\mu_1 = 0.1\mu_2$, $\mu_3 = 0.3\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$ and $a/b = 0.5$ characterized by $A + B > 0$ and $C + D > 0$

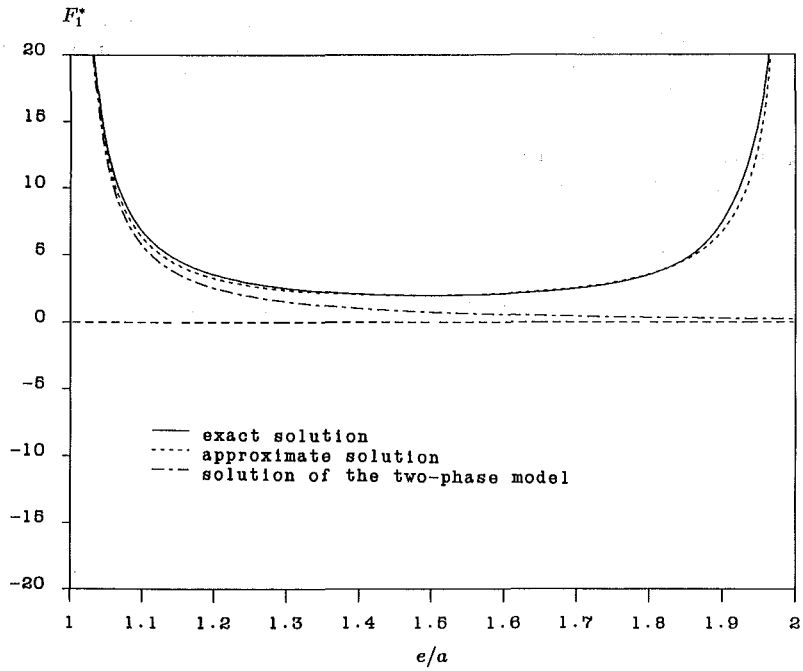


Fig. 5 Dimensionless force on the dislocation b_x in the three-phase cylinder with $\mu_1 = 4\mu_2$, $\mu_3 = 0.2\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$ and $a/b = 0.5$ characterized by $A + B < 0$ and $C + D > 0$

Based on (36), the equilibrium position of dislocation is characterized by a cubic equation

$$\beta^3 + p\beta + q = 0 \tag{37}$$

where $p = 2(A + B)/(C + D) - 1$, $q = 2(A + B)/\delta(C + D)$.

Define the dimensionless force on dislocation as $F_1^* = \pi(1 + \kappa_2)aF_1/\mu_2 b_x^2$, $F_2^* = \pi(1 + \kappa_2)aF_2/\mu_2 b_y^2$. From (36) we have

$$\frac{\partial F_1^*}{\partial \beta} = \frac{\partial F_2^*}{\partial \beta} = (A+B)\,\frac{3\beta^2 - 1}{\beta^2(\beta^2-1)^2} + (C+D)\,\frac{2\delta^2}{(1-\beta\delta)^2}. \tag{38}$$

Thus, when both the conditions $A + B < 0$ and $C + D < 0$ are satisfied, in the interval $1 < \beta < 1/\delta$ we have $\partial F_1^*/\partial\beta = \partial F_2^*/\partial\beta < 0$ which means F_1^* and F_2^* monotonically decrease. Therefore, this first-order approximation shows there is only one equilibrium position in the interval. Similarly, when $A + B > 0$ and $C + D > 0$, the approximation predicts that there is only one unstable equilibrium position in the interval.

From the foregoing discussion, it is also observed that in the three-phase model the stability behavior of a dislocation is basically independent of the orientation of its Burgers vector. This prediction is pronouncedly different from that based on the two-phase model (Dundurs and Mura, 1964).

From Figs. 3 to 6 we illustrate the variation of the values

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

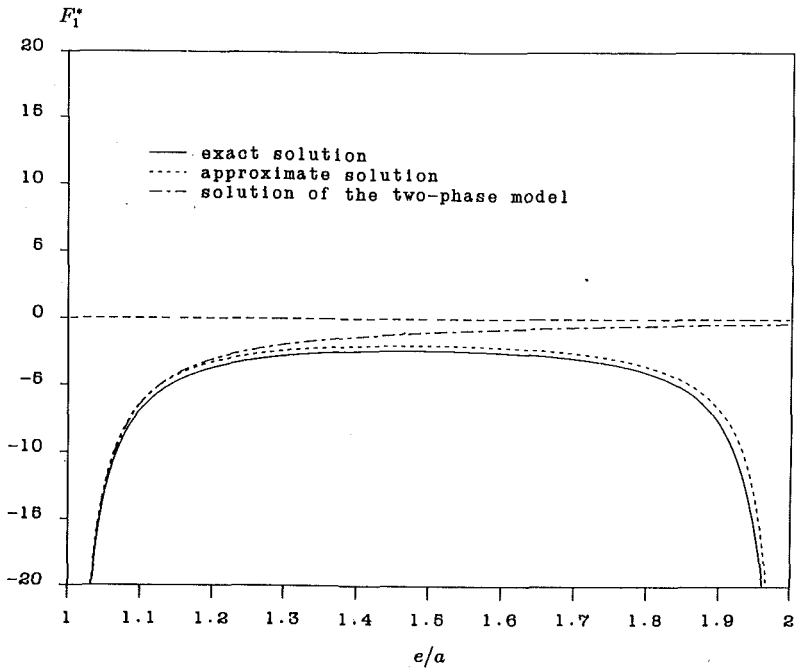


Fig. 6 Dimensionless force on the dislocation b_x in the three-phase cylinder with $\mu_1 = 0.2\mu_2$, $\mu_3 = 4\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$ and $a/b = 0.5$ characterized by $A + B > 0$ and $C + D < 0$

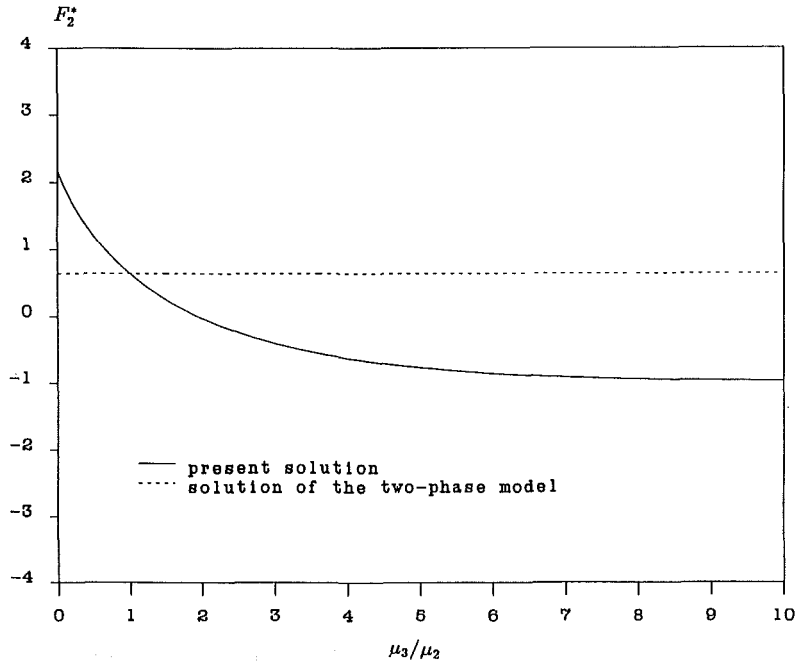


Fig. 7 Dimensionless force on the dislocation b_y in the three-phase cylinder with $\mu_1 = 10\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$, $e/a = 1.5$ and $a/b = 0.5$

of F_1^* with respect to the parameter $\beta = e/a$ for the selected four particular materials which are characterized by $A + B < 0$ and $C + D < 0$, $A + B > 0$ and $C + D > 0$, $A + B < 0$ and $C + D > 0$, and $A + B > 0$ and $C + D < 0$, respectively. Both the exact solutions calculated from equations (31) and (33) and the approximate solutions calculated from (36) are displayed. It is seen that the first-order approximation gives fairly good results for engineering purposes. For comparison, the corresponding results of the two-phase model (Dundurs and Mura, 1964) are also given in the figures. The three-phase model solution shows stable equilibrium position in Fig. 3 and unstable equilibrium position in Fig. 4, while the two-phase model solution does not. In Figs. 5 and 6 neither the three-phase model solutions nor the two-phase model solutions present any equilibrium position within the whole interval.

It is of theoretical interest to examine the influence of the elastic property of the outer phase and the geometric parameters of the body on the force on dislocation. In Fig. 7 we plot F_2^* versus $l = \mu_3/\mu_2$ for a solid consisting of three different isotropic phases. It is shown that when the rigidity of the outer phase is low, the dislocation tends to be repelled by harder inclusion; when the rigidity of the outer phase become hard, the dislocation will be repelled by the outer phase. In Fig. 8 F_2^* is plotted versus $1/\delta = b/a$ for a solid made of three dissimilar isotropic phases. When the value of b/a increases,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

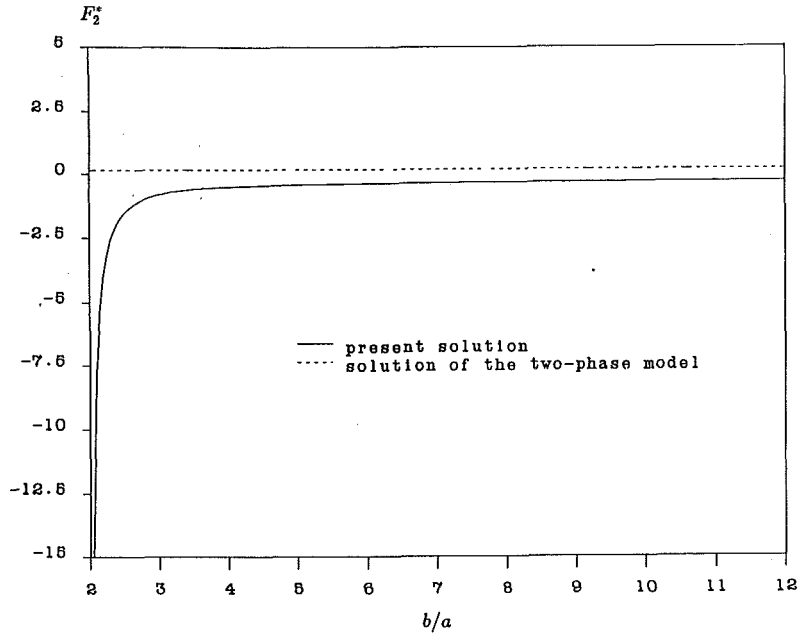


Fig. 8 Dimensionless force on the dislocation b_y in the three-phase cylinder with $\mu_1 = 10\mu_2$, $\mu_3 = 8\mu_2$, $\nu_1 = \nu_2 = \nu_3 = 0.3$ and $e/a = 1.5$

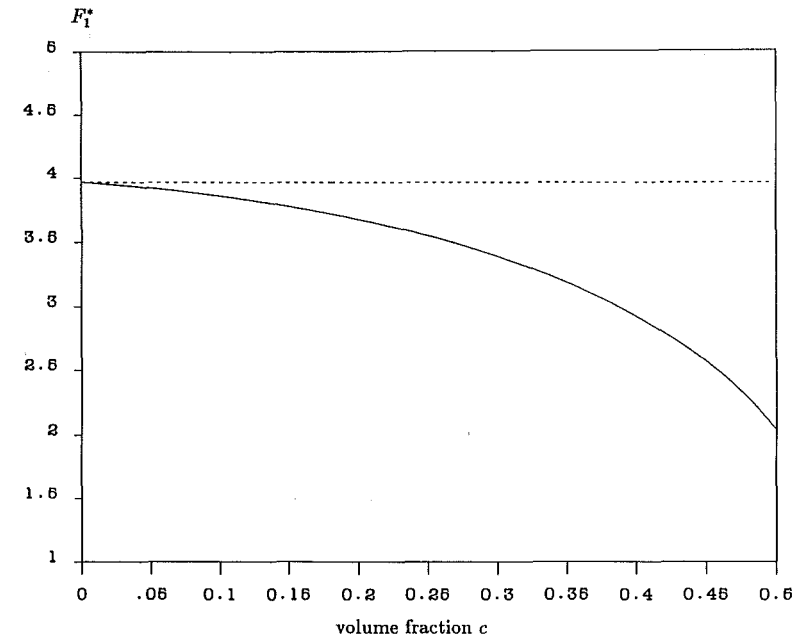


Fig. 9 Dimensionless force on the dislocation b_x for the two-phase composite with $\mu_1 = 23\mu_2$, $\nu_1 = 0.3$, $\nu_2 = 0.35$

the solutions given by this theory slowly approach to that derived from the two-phase model.

In Fig. 9 we illustrate the variation of the force on dislocation versus the volume fraction c for a two-phase composite with $m = \mu_1/\mu_2 = 23$, $\nu_1 = 0.3$ and $\nu_2 = 0.35$ where the elastic constants of the composite, namely, the outer phase are determined by means of the modified Mori-Tanaka method suggested by Luo and Weng (1989). It is observed that in this case the force decreases with the increase of the concentration of fibers which are harder than the matrix.

In the above calculations, we truncate the Laurent series at $n = \pm 20$. The results have indicated very good convergence of the series form solutions.

References

Christensen, R. M., and Lo, K. H., 1979, "Solution for Effective Shear Properties in Three Phase Sphere and Cylinder Models," *Journal of Mechanics and Physics of Solids*, Vol. 27, pp. 315–330.

Dundurs, J., 1967, "Elastic Interaction of Dislocations with Inhomogeneities," *Mathematical Theory of Dislocations*, T. Mura, ed., ASME, New York, pp. 70–115.

Dundurs, J., and Mura, T., 1964, "Interaction Between an Edge Dislocation and a Circular Inclusion," *Journal of the Mechanics and Physics of Solids*, Vol. 12, pp. 177–189.

England, A. H., 1971, *Complex Variable Methods in Elasticity*, John Wiley and Sons, New York.

Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, "Interaction Between a Circular Inclusion and an Arbitrarily Oriented Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 1007–1013.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Head, A. K., 1953a, "Edge Dislocations in Inhomogeneous Media," *Proceedings of the Physical Society of London*, Vol. B66, pp. 793–801.

Head, A. K., 1953b, "The Interaction of Dislocations and Boundaries," *Philosophical Magazine*, Vol. 44, pp. 92–94.

Luo, A. H., and Weng, G. J., 1989, "On Eshelby's S-tensor in a Three-Phase Cylindrically Concentric Solid, and the Elastic Moduli of Fiber-Reinforced Composites," *Mechanics of Materials*, Vol. 8, pp. 77–88.

Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, P. Noordhoff, Groningen.

APPENDIX

$$h^{(2)}_{xx1} = -2\left(1+\frac{2x_1^2}{r_1^2}\right)\frac{y}{r_1^2}+\left[B+A\left(1+\frac{4x_2^2}{r_2^2}\right)\right]\frac{y}{r_2^2}$$
$$-\left[B+A\left(1+\frac{4x^2}{r^2}\right)\right]\frac{y}{r^2}-2\,\frac{B-A}{\beta}\,\frac{axy}{r^4}-2A\left(1-\frac{4x^2}{r^2}\right)\frac{a^2y}{r^4}$$
$$-A\,\frac{\beta^2-1}{\beta^3}\left[\left(1-\frac{4x_2^2}{r_2^2}\right)\frac{2x_2}{r_2^2}-\frac{\beta^2-1}{\beta}\left(1-\frac{4x_2^2}{r_2^2}\right)\frac{a}{r_2^2}\right]\frac{2ay}{r_2^2}$$
$$-\frac{2}{a}\sum_{n=1}^{\infty}a'_n\left(\frac{r}{a}\right)^n\sin n\theta+\frac{1}{a}\sum_{n=1}^{\infty}a'_n n\left(\frac{r}{a}\right)^n\sin(n-2)\theta$$
$$-\frac{1}{a}\sum_{n=1}^{\infty}[a'_n(n-1)+a'_{-n}]\left(\frac{r}{b}\right)^{n-2}\sin(n-2)\theta$$
$$+\frac{2A}{a}\sum_{n=1}^{\infty}a'_{-n}\left(\frac{a}{r}\right)^n\sin n\theta+\frac{A}{a}\sum_{n=1}^{\infty}a'_{-n}n\left(\frac{a}{r}\right)^n\sin(n+2)\theta$$
$$+\frac{1}{a}\sum_{n=1}^{\infty}[Ba'_n-Aa'_{-n}]\left(\frac{a}{r}\right)^{n+2}\sin(n+2)\theta$$

$$h^{(2)}_{yy1} = -2\left(1-\frac{2x_1^2}{r^2}\right)\frac{y}{r_1^2}+\left(3A-B-4A\,\frac{x_2^2}{r_2^2}\right)\frac{y}{r_2^2}$$
$$-\left(3A-B-4A\,\frac{x^2}{r^2}\right)\frac{y}{r^2}$$
$$-2A\,\frac{\beta^2-1}{\beta^3}\,\frac{a}{r_2^2}\left[\frac{6x_2y}{r_2^2}+\frac{8x_2^3y}{r_2^4}+\frac{\beta^2-1}{\beta}\left(1-\frac{4x_2^2}{r_2^2}\right)\frac{ay}{r_2^2}\right]$$
$$+\frac{\mathrm{B-A}}{\beta}\,\frac{2axy}{y^4}+2A\left(1-\frac{4x^2}{r^2}\right)\frac{a^2y}{r^4}-\frac{2}{a}\sum_{n=1}^{\infty}a'_n\left(\frac{r}{a}\right)^n\sin n\theta$$
$$-\frac{1}{a}\sum_{n=1}^{\infty}a'_n n\left(\frac{r}{a}\right)^n\sin(n-2)\theta$$
$$+\frac{1}{a}\sum_{n=1}^{\infty}[a'_n(n-1)+a'_{-n}]\left(\frac{r}{b}\right)^{n-2}\sin(n-2)\theta$$
$$+\frac{2A}{a}\sum_{n=1}^{\infty}a'_{-n}\left(\frac{a}{r}\right)^n\sin n\theta$$
$$-\frac{A}{a}\sum_{n=1}^{\infty}a'_{-n}n\left(\frac{a}{r}\right)^n\sin(n+2)\theta-\frac{1}{a}\sum_{n=1}^{\infty}[Ba'_n$$
$$-Aa'_{-n}]\left(\frac{a}{r}\right)^{n+2}\sin(n+2)\theta$$

$$h^{(2)}_{xy1} = -2\left(1-\frac{2x_1^2}{r_1^2}+(3A-B-4A\,\frac{x_2^2}{r_2^2}\right)\frac{x_2}{r_2^2}$$
$$-2A\,\frac{\beta^2-1}{\beta^3}\left[1-\frac{8x_2^2}{r_2^2}+\frac{8x_2^4}{r_2^4}\right.$$
$$+\frac{\beta^2-1}{\beta}\left(3-4\,\frac{x_2^2}{r_2^2}\right)\frac{ax_2}{r_2^2}\Big]\frac{a}{r_2^2}-\left(3A-B-4A\,\frac{x^2}{r^2}\right)\frac{x}{r^2}$$
$$-\frac{B-A}{\beta}\left(1-2\,\frac{x^2}{r^2}\right)\frac{a}{r^2}+2A\left(3-4\,\frac{x^2}{r^2}\right.$$
$$+\frac{1}{a}\sum_{n=1}^{\infty}a'_n n\left(\frac{r}{a}\right)^n\cos(n-2)\theta$$
$$-\frac{1}{a}\sum_{n=1}^{\infty}[a'_n(n-1)+a'_{-n}]\left(\frac{r}{a}\right)^{n-2}\cos(n-2)\theta$$
$$-\frac{A}{a}\sum_{n=1}^{\infty}a'_{-n}n\left(\frac{a}{r}\right)^n\cos(n+2)\theta$$
$$+\frac{1}{a}\sum_{n=1}^{\infty}[Aa'_{-n}(n+1)$$
$$-Ba'_n]\left(\frac{a}{r}\right)^{n+2}\cos(n+2)\theta$$

$$h^{(2)}_{xx2} = -2\left(1-\frac{2x_1^2}{r_1^2}\right)\frac{x_1}{r_1^2}+\left[B+A\left(1-\frac{4x_2^2}{r_2^2}\right)\right]\frac{x_2}{r_2^2}$$
$$-\left[B+A\left(1-\frac{4x^2}{r^2}\right)\right]\frac{x}{r^2}$$
$$-A\,\frac{\beta^2-1}{\beta^3}\left[2\beta^2\left(1-\frac{2x_2^2}{r_2^2}\right)+4\left(3-\frac{4x_2^2}{r_2^2}\right)\frac{x_2^2}{r_2^2}\right.$$
$$-\frac{2(\beta^2-1)}{\beta}\left(3-4\,\frac{x_2^2}{r_2^2}\right)\frac{ax_2}{r_2^2}\Big]\frac{a}{r_2^2}+2A\left(3-\frac{4x^2}{r^2}\right)\frac{a^2x}{r^4}$$
$$+\frac{1}{\beta}\left(1-\frac{2x^2}{r^2}\right)(2A\beta^2+B-A-2N)\frac{a}{r^2}$$
$$-2Na_0\left(1-\frac{2x^2}{r^2}\right)\frac{a}{r^2}+\frac{2}{a}\sum_{n=1}^{\infty}a_n\left(\frac{r}{a}\right)^n\cos n\theta$$
$$-\frac{1}{a}\sum_{n=1}^{\infty}a_n n\left(\frac{r}{a}\right)^n\cos(n-2)\theta$$
$$+\frac{1}{a}\sum_{n=1}^{\infty}[a_n(n-1)a_{-n}]\left(\frac{r}{a}\right)^{n-2}\cos(n-2)\theta$$
$$+\frac{2A}{a}\sum_{n=1}^{\infty}a_{-n}\left(\frac{a}{r}\right)^n\cos n\theta$$
$$+\frac{A}{a}\sum_{n=1}^{\infty}a_{-n}n\left(\frac{a}{r}\right)^n\cos(n+2)\theta$$
$$-\frac{1}{a}\sum_{n=1}^{\infty}[Ba_n+A(n+1)a_{-n}]\left(\frac{a}{r}\right)^{n+2}\cos(n+2)\theta$$

$$h^{(2)}_{yy2} = 2\left(3-\frac{2x_1^2}{r_1^2}\right)\frac{x_2}{r_1^2}-\left(5A+B-4A\,\frac{x_2^2}{r_2^2}\right)\frac{x_2}{r_2^2}$$
$$+\left(5A+B-4A\,\frac{x^2}{r^2}\right)\frac{x}{r^2}-A\,\frac{\beta^2-1}{\beta^3}\left[2(2-\beta^2)-4(5-\beta^2)\,\frac{x_2^2}{r_2^2}\right.$$
$$+16\,\frac{x_2^4}{r_2^4}+\frac{2(\beta^2-1)}{\beta}\left(3-4\,\frac{x_2^2}{r_2^2}\right)\frac{ax_2}{r_2^2}\Big]\frac{a}{r_2^2}$$
$$-2A\left(3-\frac{4x^2}{r^2}\right)\frac{a^2x}{r^4}-\frac{1}{\beta}\left(1-\frac{2x^2}{r^2}\right)(2A\beta^2+B$$
$$-A-2N)\,\frac{a}{r^2}-2Na_0\left(1-\frac{2x^2}{r^2}\right)\frac{a}{r^2}+\frac{2}{a}\sum_{n=1}^{\infty}a_n\left(\frac{r}{a}\right)^n\cos n\theta$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+\frac{1}{a}\sum_{n=1}^{\infty} a_n n\left(\frac{r}{a}\right)^n \cos(n-2)\theta$$

$$-\frac{1}{a}\sum_{n=1}^{\infty}[a_n(n-1)-a_{-n}]\left(\frac{r}{a}\right)^{n-2}\cos(n-2)\theta$$

$$+\frac{2A}{a}\sum_{n=1}^{\infty} a_{-n}\left(\frac{a}{r}\right)^n \cos\, n\theta - \frac{A}{a}\sum_{n=1}^{\infty} a_{-n} n\left(\frac{a}{r}\right)^n \cos(n+2)\theta^{\,n}$$

$$+\frac{1}{a}\sum_{n=1}^{\infty}[Ba_n + A(n+1)a_{-n}]\left(\frac{a}{r}\right)^{n+2}\cos(n+2)\theta \qquad 2)\theta$$

$$h^{(2)}_{xy2} = -2\left(1-\frac{2x_1^2}{r_1^2}\right)\frac{y}{r_1^2} + \left[B + A\left(1-\frac{4x_2^2}{r_2^2}\right)\right]\frac{y}{r_2^2} - \left[B + A\left(1-\frac{4x^2}{r^2}\right)\right]\frac{y}{r^2}$$

$$+2A\,\frac{\beta^2-1}{\beta^3}\left[\left(2\beta^2-4+8\,\frac{x_2^2}{r_2^2}\right)\frac{x_2 y}{r_2^2} + \frac{\beta^2-1}{\beta}\left(1-\frac{4x_2^2}{r_2^2}\right)\frac{ay}{r_2^2}\right]\frac{a}{r_2^2}$$

$$+2A\left(1-\frac{4x^2}{r^2}\right)\frac{a^2 y}{r^4} - \frac{1}{\beta}(2A\beta^2 + B - A - 2N)\,\frac{2axy}{r^4} - 4Na_0\,\frac{axy}{r^4}$$

$$+\frac{1}{a}\sum_{n=1}^{\infty} a_n n\left(\frac{r}{a}\right)^n \sin(n-2)\theta$$

$$-\frac{1}{a}\sum_{n=1}^{\infty}[a_n(n-1)-a_{-n}]\left(\frac{r}{a}\right)^{n-2}\sin(n-2)\theta$$

$$+\frac{A}{a}\sum_{n=1}^{\infty} a_{-n} n\left(\frac{a}{r}\right)^n \sin(n+2)\theta - \frac{1}{a}\sum_{n=1}^{\infty}[Aa_{-n}(n+1)$$

$$+Ba_n]\left(\frac{a}{r}\right)^{n+2}\sin(n+2)\theta \qquad \text{(A1)}$$

where

$$x_1 = x-e, \quad x_2 = x-\frac{a^2}{e},$$

$$r_1^2 = (x-e)^2+y^2, \quad r_2^2 = \left(x-\frac{a^2}{e}\right)^2+y^2. \qquad \text{(A2)}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

On the Extent of Dominance of Asymptotic Elastodynamic Crack-Tip Fields: Part I—An Experimental Study Using Bifocal Caustics

Sridhar Krishnaswamy
Research Fellow

Ares J. Rosakis
Assoc. Professor,
Assoc. Mem. ASME

Graduate Aeronautical Laboratories,
California Institute of Technology,
Pasadena, CA 91125

The question of the domain of dominance of mode I asymptotic elastodynamic crack-tip fields is investigated experimentally for the cases of dynamically loaded stationary cracks as well as dynamically propagating cracks. The experiments reported are on three-point bend specimens loaded dynamically using a drop-weight tower. An optical configuration leading to a bifocal *high-speed camera is proposed. This is used in conjunction with the method of caustics to obtain apparent dynamic stress intensity factor measurements simultaneously from two different regions around the crack tip. The results of this study indicate that three-dimensional and transient effects necessarily have to be taken into account in the interpretation of dynamic fracture experiments.*

1 Introduction

The mechanics of crack initiation and propagation under conditions where material inertial effects are important—either due to rapid crack propagation or due to transient loading—fall into the broad field of study of dynamic fracture mechanics. While the micromechanical aspects of the actual cleavage or void coalescence that are an essential part of the fracture process are inaccessible through continuum descriptions, much of the theoretical basis of dynamic fracture mechanics derives from continuum analyses. The attempt to characterize the initiation and dynamic propagation of a crack in a material is typically done in two parts. A fracture criterion is first postulated which would relate some parameter in a continuum solution to a material specific property. For example, a crack might be postulated to initiate and propagate when the strain at some prescribed point along the crack line reaches a critical value which is supposedly a material property determinable only through experiments. The correctness and utility of such a fracture criterion must then be established by showing its validity under different geometric and loading configurations. That is, while the details of the stress and strain fields might depend on the particular geometric and loading conditions, the critical values at which a crack would initiate and grow must be configuration independent.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 31, 1989; final revision, Mar. 14, 1990.

By far the most widely quoted fracture criterion in the dynamic fracture of nominally brittle materials—defined here as materials where no extensive plastic deformation occurs prior to or during crack growth—is that based on the stress intensity factor. This is a parameter that arises naturally from certain mathematically idealized crack problems and, since its measurability and utility as a fracture criterion form the subject of inquiry of this paper, it is probably instructive at this point to delve into its origins.

The Asymptotic Field. The first dynamic fracture problems to be analytically solved were those where the material was modeled as linear elastic and the geometry and loading were taken to conform to plane-strain conditions. The crack solutions of Yoffe (1951), Craggs (1960), Broberg (1960), Baker (1962), and Freund (1972) all fall into this category. While the problems considered by them differed in some vital respects—some were transient while others were steady-state—a common outcome of all their solutions was the nature of the asymptotic stress field. Indeed, as has been shown by Freund and Clifton (1974), the mode I stress field with reference to a Cartesian coordinate system moving with the crack tip of "all plane elastodynamic solutions for (smoothly turning) running cracks, for which the total internal energy is finite," can be asymptotically described by the square-root singular expression

$$\sigma_{\alpha\beta} = \frac{K_I^d(t)}{\sqrt{2\pi r}} \Sigma_{\alpha\beta}^I(\theta, \dot{a}) + O(1), \quad \text{as } r \to 0. \tag{1}$$

Here, (r,θ) is a polar coordinate system travelling with the crack tip, $\Sigma_{\alpha\beta}^I$ are known universal functions of θ and crack

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

velocity $\dot{a}$ (with respect to some fixed coordinate frame), and K_I^d is the so-called mode I stress intensity factor dependent only on the specific geometric and loading conditions of a problem. The above will henceforth be referred to as the dynamic *stress intensity factor field* or, more concisely, the K_I^d-field. Note that this is strictly asymptotic and that any attempt to use it over a finite extent must be justified by showing that the asymptotic solution dominates (to within some acceptable error) over this region. A region where the above field dominates will then be referred to as a region of K_I^d-*dominance.*

The situation of a stationary crack that is loaded dynamically is less clear but the few analytical solutions that exist—Achenbach (1973)—do exhibit a stress intensity factor field as given in equation (1) with $\dot{a} \to 0$. While the universality of such a field has not been rigorously established for all stationary cases, the leeway provided by the unspecified extent of validity of the asymptotic field has allowed for the assumption that the stress intensity factor field is descriptive of all plane elastodynamic crack problems.

It has of course been recognized that the asymptotic field as given above leads to singular stresses at the crack tip and, as such, plasticity and other nonlinear effects must come into play near the crack tip. The possible relevance of the stress intensity factor field thus arises from an argument of confined plasticity: If the length scale over which plasticity effects are important is much smaller than the length over which the asymptotic K_I^d-field is dominant, then the stress intensity factor field must characterize the plastic and process zones and hence the fracture process itself. This is the so-called small-scale yielding concept.

Finally, in an attempt to model the more realistic problem of fracture in thin plates, the asymptotic field for conditions of plane stress are formally obtained by a suitable modification of the material parameters—"formal" in the sense that no attempt is made to inquire into the acceptability of the plane-stress approximation in the vicinity of a crack tip.

Fracture Criteria. The presumed universality of the stress intensity factor field for problems of small-scale yielding provides a possible fracture criterion of the form

$$K_I^d(t) = K_{ID}. \tag{2}$$

Here, the left side is the stress intensity factor, in principle entirely determined through an analysis of a boundary-initial value problem, and the right side represents a material quantity called the dynamic fracture toughness that can only be determined through experiments. The dynamic stress intensity factor $K_I^d(t)$ is a determinable function of crack length $a(t)$, crack velocity $\dot{a}(t)$, some generalized measure of the applied load $P(t)$, and geometry. Also, based on perceived changes in the resistance to crack propagation with changes in rate of loading, it has been suggested that the dynamic fracture toughness must be dependent on crack velocity. Thus, the fracture criterion as given in equation (2) becomes

$$K_I^d(a(t), \dot{a}(t), P(t)) = K_{ID}(\dot{a}(t), \ldots), \tag{3}$$

which is then an evolution equation for crack growth (equation of motion for the crack tip) and therein lies the attractiveness of equation (2) as a fracture criterion.

The primary focus of dynamic fracture mechanics has been the determination of the fracture toughness and its possible dependence on parameters such as crack velocity, acceleration, and temperature. The extensive experimental effort that has gone into this task over the past two decades has generated a large amount of data without conclusively deciding on the appropriateness of equation (2) as a fracture criterion. For the most part, experimental techniques which *assume* the dominance of the asymptotic K_I^d-field over some region around the crack tip have been utilized to argue about the function dependence of the dynamic fracture toughness on crack velocity, acceleration, etc. Using this approach, some authors—Kobayashi and Dally (1977), Bilek (1980), Rosakis, Duffy, and Freund (1984), and Zehnder and Rosakis (1989)—have found that the dynamic fracture toughness depends in a one-to-one manner on the crack velocity while others—Kobayashi and Mall (1978) and Ravi-Chandar and Knauss (1984)—do not find support for such a relation. Takahashi and Arakawa (1987) have argued for an acceleration dependence of the dynamic fracture toughness while Eshelby (1969) has pointed out that "a crack tip exhibits no inertia." Dahlberg, Nilsson, and Brickstad (1980) and Kalthoff (1983) have experimentally found that the dynamic fracture toughness has a systematic specimen dependence. If correct, this would imply that the stress intensity factor based fracture criterion is flawed. The reasons for these disagreements have been widely debated but a reasoned consensus is yet to emerge. For the most part, the variations have been attributed to differences in material response.

Of late, however, attention has been increasingly drawn toward another possible source of divergence, namely that the interpretation of the experimental data has, to a large extent, required the assumption that the asymptotic field be valid over at least the region over which the measurements were made. Realizing that this might be a source of error, Dally, Fourney, and Irwin (1985) have developed "higher-order solutions" to the two-dimensional, dynamic, steady crack propagation problem to aid in the interpretation of their experimental photoelastic fringes. However, questions regarding the validity of these solutions and the errors associated with time-averaging the field over large domains, at least insofar as truly transient problems are concerned, have been raised by Knauss and Ravi-Chandar (1985). While Dally, Fourney, and Irwin (1985) are probably correct in maintaining that the asymptotic field might not be adequate for transient problems, it is precisely these problems that are also not amenable to the kind of steady-state higher-order solutions that they propose.

For stress-wave loaded configurations, the time taken to fully establish a K_I^d-field at some distance from the crack tip is an issue of great importance. This has been studied experimentally by Ravi-Chandar and Knauss (1987) and analytically by Ma and Freund (1986). One interesting conclusion from these studies is that, even in a purely two-dimensional setting, the time scale for establishment of K_I^d-dominance at distances sufficiently far from the crack-tip, as to be accessible for experimental scrutiny, is of the same order as other relevant time scales in the problem. This raises the possibility that K_I^d-dominance in these stress-wave loaded configurations could be confined to a very small region around the crack-tip. In a real laboratory specimen, however, strong three-dimensional effects are expected to prevail near the crack front.

Among the first to study the question of crack-tip three-dimensionality were Rosakis and Ravi-Chandar (1986). They experimentally sought to determine the extent to which three-dimensionality affected measurements made using the method of caustics under *static conditions.* Their findings indicate that three-dimensional effects are confined to a radial distance of one-half plate thickness around the crack tip beyond which they find the two-dimensional static stress intensity factor field prevailing over some distance. Supporting analytical basis for this work was provided by Yang and Freund (1985) who again find that the three-dimensional character of the field is confined to about one-half of the plate thickness. Many other numerical studies (Parsons, Hall, and Rosakis, 1986; Nakamura and Parks, 1987) also support such a conclusion at least for elastostatic conditions. Smith and Freund (1988) have recently reported similar results for the elastodynamic steady crack propagation problem. The numerical analyses of Nakamura and Parks (1987) and Smith

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

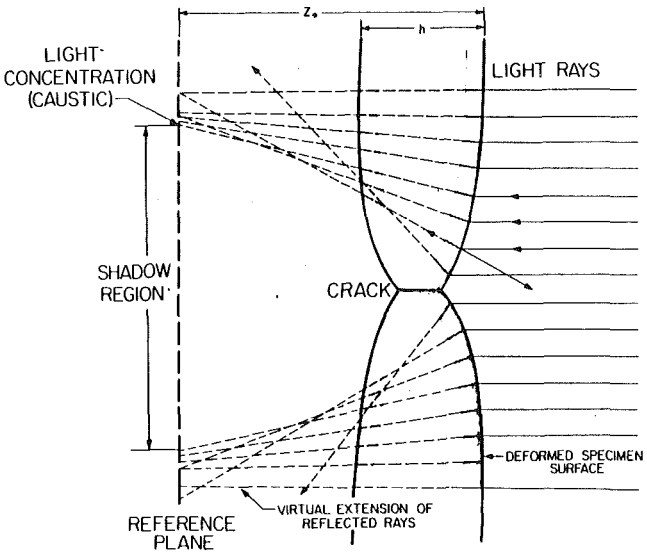


Fig. 1(a) Schematic diagram of the method of caustics in reflection

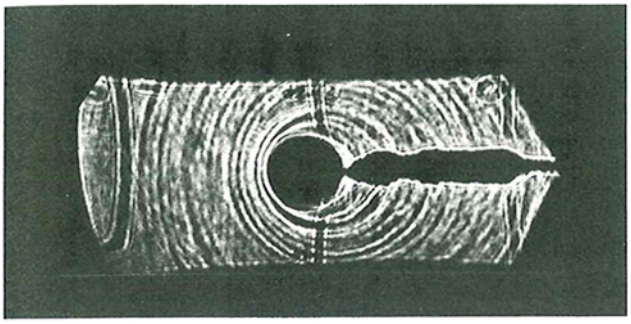


Fig. 1(b) Example of reflected caustic from a propagating crack tip

and Freund (1988), however, *impose* the plane-stress asymptotic field at some arbitrary distance from the crack front. These studies, therefore, find that outside of the three-dimensional zone, the asymptotic K_I-field is recovered. Too much importance, however, should not be attached to this latter result in view of the fact that this outcome was most definitely influenced by their having *prescribed* the K_I-field at some *definite* distance *greater* than one-half of plate thickness.

Apart from the works mentioned above, many analytical and numerical studies have attempted to incorporate nonlinearity of material response. Among these, Lam and Freund (1985) and Freund and Hutchinson (1985) adopt the small-scale yielding approach; i.e., it is assumed that a K-dominant field surrounds the nonlinear zone and the details of the continuum structure in the plastic region are obtained in this context. In particular, Lam and Freund (1985) have sought to provide an analytical basis for the uniqueness of the K_{ID}—$\dot{a}$ relation by postulating a more fundamental fracture criterion inside the plastic zone and then studying its ramifications for the surrounding K_I^d-field. Essentially the same approach is taken by Freund and Hutchinson (1985) where a material exhibiting rate-dependent plastic behavior is considered. For the purposes of this paper, the point to note is that the common thread underlying all these studies is the acceptance of a K-dominant field over some annulus around the crack tip.

Overview of Paper. This paper documents an effort to systematically investigate the validity and extent of dominance of the stress intensity factor field under conditions of dynamic loading and propagation. The question that this work addresses is the following: *Is there a region around the crack-tip where the stress and displacement fields could be said to be modeled by the two-dimensional dynamic stress intensity factor field?* This will be answered within the context of a three-point bend specimen loaded dynamically in a drop-weight tower—an experimental configuration that has been extensively used in the past by Kalthoff (1985) and Zehnder and Rosakis (1989). The findings of this work are specific to this particular configuration. The questions raised here and the methodology adopted in trying to answer them, of course, have a wider applicability.

2 The Method of Caustics

A detailed exposition of the history and theory of the method of caustics can be found in Manogg (1964), Beinert and Kalthoff (1981), and Rosakis and Zehnder (1985). Here only a brief outline of the method of caustics in reflection is presented. A schematic diagram of the experimental method is shown in Fig. 1(*a*). A collimated laser beam is incident on a polished, optically flat fracture specimen and the reflected light is collected and imaged in a camera. The camera is set to focus on a virtual object plane at a distance z_0 behind the specimen. When the specimen is loaded, its deformed surface could be described by $x_3 = -f(x_1, x_2)$, where (x_1, x_2) are points on the specimen. Due to this deformation, a light ray incident at a point (x_1, x_2) on the specimen gets mapped to a point (X_1, X_2) on the virtual object plane of the camera. The mapping can be shown to be given by

$$\mathbf{X} \simeq \mathbf{x} - 2z_0 \nabla f \tag{4}$$

where it has been assumed that $z_0 \gg \max |f(x_1, x_2)|$ which is true for all experiments reported in this paper.

Caustics in Elastodynamic Fracture. Consider a fracture specimen made of a linear elastic, homogeneous, isotropic solid subject to mode I, generalized plane-stress conditions. Suppose that the specimen is dynamically loaded or that the crack is dynamically propagating. Contingent on the assumption that there is a region near the crack tip where the stress field is given by equation (1), it can be shown that the caustic map of the out-of-plane displacement field corresponding to (1) leads to a characteristic shadow spot—a dark region where no light rays get mapped, surrounded by a bright epicycloid called the caustic (see Fig. 1(*b*)). The transverse diameter D of the caustic (which, in general, varies with time) can be related (Rosakis, Duffy, and Freund, 1984) to the instantaneous stress intensity factor through

$$K_I^d = \frac{ED^{5/2}}{10.7 z_0 \nu h} F(\dot{a}) C(\alpha_l) \tag{5}$$

where

$$F(\dot{a}) = \frac{4\alpha_l \alpha_s - (1+\alpha_s^2)^2}{(1+\alpha_s^2)(\alpha_l^2 - \alpha_s^2)} \tag{6}$$

and

$$C(\alpha_l) = (6.8 + 14.4\alpha_l - 2.6\alpha_l^2)/18.6 \tag{7}$$

and

$$\alpha_{l,s} = (1 - (\dot{a}/c_{l,s})^2)^{1/2} \tag{8}$$

where c_l and c_s are the longitudinal and shear wave speeds of the material. The locus of points on the specimen surface that map on to the caustic curve is called the initial curve. For sufficiently small crack velocities ($\dot{a} < 0.3c_s$), the initial curve is very nearly circular and its instantaneous nominal radius can be approximated by

$$r_0(t) = \left(\frac{3h\nu K_I^d(t) z_0}{2\sqrt{2\pi}E}\right)^{2/5} F(\dot{a})^{(-2/5)}. \tag{9}$$

The above equations in the limit $\dot{a} \to 0$ are valid for the case of a dynamically loaded stationary crack. In a typical experiment, a high-speed camera would be used to obtain a time se-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

quence of caustics from whose diameter $D(t)$, the stress intensity factor $K_I^d(t)$, would be computed using (5). From equation (9), it is seen that $r_0(t) = \hat{r}_0(K_I^d(t), z_0, \dot{a})$ and as $K_I^d(t)$ varies with time in a dynamic experiment, the radius of the initial curve would perforce change during the course of the experiment. It is thus vital to know that equation (1) leads to a valid expression for the displacement field for at least the range of radii that the initial curve would cover during the event of interest.

To make matters a bit more explicit, suppose that during the event of interest (which could be the crack propagation phase of a test) the initial curve radius is known from previous experience to vary in the region $r_{\min} \le r_0 \le r_{\max}$ for some choice of the object plane distance z_0. Also, since noncontinuum, nonlinear, and three-dimensional effects near the crack tip preclude the application of (1) in the limit $r \to 0$, suppose that the asymptotic field as given by (1) is actually a good approximation to reality in an annulus surrounding the crack tip as shown in Fig. 2. Thus,

$$\sigma_{\alpha\beta} \approx \frac{K_I^d(t)}{\sqrt{2\pi r}} \Sigma_{\alpha\beta}(\theta, \dot{a}), \quad r_{\text{in}} < r < r_{\text{out}}. \tag{10}$$

Since the domain of validity of (10) could in general be time dependent as well, let r_{in} and r_{out} be such as to give the smallest annulus in which (10) holds for all times during the entire event of interest. For a valid interpretation of the caustics, the inequalities

$$r_{\text{in}} < r_{\min} \le r_0 \le r_{\max} < r_{\text{out}} \tag{11}$$

must be satisfied in order to have the initial curve fall in a region of K_I^d-dominance.

The first attempts to address the issue of the validity of (10) are reported in Ravi-Chandar and Knauss (1987). A series of tests was performed using the method of caustics in transmission on identical specimens under identical stress-wave loading, varying from test to test only the object plane distance z_0. In this manner a range of initial curve radii was scanned and since, presumably, the actual stress intensity factor history $K_I^d(t)$ for the various tests must be identical, the apparent stress intensity factors measured from caustics obtained from different object planes must also agree, at least for those times when the requirements of (11) are met. Their findings indicate that the assumption behind (10) might not be tenable, at least for their experiments. Since a substantial part of the dynamic fracture data extent in the literature has been obtained through the use of either photo-elasticity or the method of caustics, the ramifications of this result are potentially far reaching and thus deserve greater scrutiny.

In particular, it would be preferable to be able to obtain the apparent stress intensity factor values from different initial curves around the crack tip for the same specimen at any given instant in time. While comparison of experimental results across tests can be made with confidence for the loading regime of the experiments, it becomes more difficult in the crack propagating phase. This is because small variations in the crack motion history could lead, as will be seen shortly, to large variations in the values for the measured stress intensity factor. It is in an attempt to avoid this problem that one would like to be able to obtain caustics from different initial curves simultaneously from the same experiment. Conceptually, to set up an experiment to do this is an easy matter. The light reflected off a dynamically loaded fracture specimen could be split into many paths (using beam splitters) which could then be piped into a series of cameras each of which is set to focus on a different object plane. The resulting caustics imaged in these cameras would have been obtained from different initial curves at each instant of time. Here, however, an alternate scheme is proposed that could, in fact, be incorporated into essentially any high-speed photography system. An optical configuration that would allow for the *simultaneous* acquisition of dynamic caustics from *two* different object planes using only a *single* high-speed camera will now be discussed.

3 The Bifocal High-Speed Camera

In the previous section it was noted that the initial curve is the locus of points on the specimen which optically map onto the caustic, and thus it is the region from which the apparent stress intensity factor value is measured. Also, for a given stress intensity factor value, it is the distance between the specimen and the camera (virtual) object plane that determines the radius of the initial curve. Thus, if two different object planes can be simultaneously brought into focus in a single camera, stress intensity factor information from two different initial curves would be obtained at the same time. Figure 3 shows a schematic of the optical set-up required to bring two

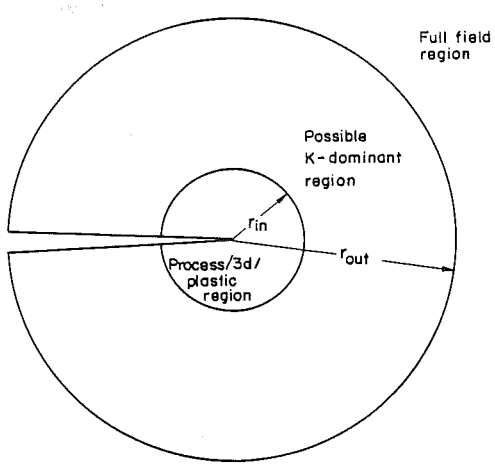


Fig. 2 Possible region of K_I^d-dominance around the crack tip

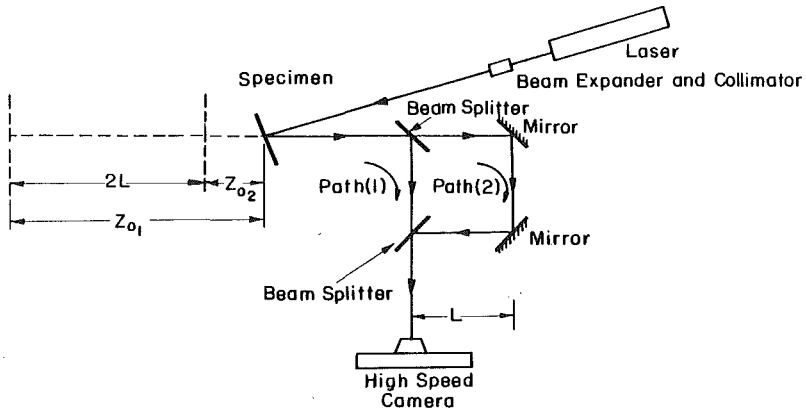


Fig. 3 Schematic diagram of the bifocal set-up

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

different object planes simultaneously into focus in a single camera. The set-up entails the use of two beam splitters and two mirrors by which two optical paths of different path lengths are established between the specimen and the high-speed camera. With reference to the figure, let the high-speed camera be set up to focus at a distance g from the camera lens. Along path (1) this would mean that the virtual object plane is located at a distance z_{01} behind the specimen. Along path (2), however, by virtue of the increased length ($2L$) between the specimen and the camera lens, the object plane would now fall only at a distance $z_{02} = z_{01} - 2L$ behind the specimen. Thus, the caustics obtained from the two paths would be from two different initial curve radii. They could be made to appear on the film track of the high-speed camera either superposed or side by side.

For nomenclatural convenience, the pairs of dynamic caustics obtained by use of this bifocal high-speed camera will henceforth be referred to as "bifocal caustics," with the implicit understanding that such caustics are obtained from two different initial curves on the same specimen at the same time. (There is, of course, a time delay of $2L/c$ because of the finiteness of the speed of light c, but this is negligible for values of L used in practice.) By changing the distance L from test to test or by changing the focal plane of the camera, one could, of course, scan various sets of initial curves. The comparison of whether the apparent stress intensity factors obtained from the bifocal caustics agree over some range of r_0 would, however, be done solely within each test for reasons stated earlier.

Since the largest measurement uncertainties are in the crack velocities, one could check for the following ratio (as obtained from equation (5))

$$\frac{K_{I1}^d}{K_{I2}^d} = \left(\frac{z_{02}}{z_{01}}\right)\left(\frac{D_1}{D_2}\right)^{5/2} \tag{12}$$

where subscript "1" refers to quantities corresponding to path (1) and subscript "2" to path (2). This ratio should be unity if the initial curve corresponding to z_{01} and z_{02} both fall within the K_I^d-dominant region. Any substantial deviation from unity of this ratio is then an indication of a lack of K_I^d-dominance over the corresponding region. The above ratio has the added advantage that the picture magnification and plate thickness get scaled out as well. Thus, the measurement uncertainty in evaluating the above ratio of apparent stress intensity factors through (12) is less than in the evaluation of the individual apparent stress intensity factors through (5).

4 Experiment

The experiments reported in this work were designed to measure: (a) the apparent stress intensity factor history using the bifocal caustics technique described in the previous section and (b) the time history of the traction boundary conditions that the specimen experiences for subsequent input into a dynamic finite element analysis. In this paper, only the findings of the bifocal caustics experiments will be described. The numerical simulation of the drop-weight tests are the subject of Part II of this paper.

Specimen Preparation. The specimens were made of either plain or vacuum-arc remelt AISI 4340 steel. After preliminary machining, the specimens were heat treated as follows: (1) normalized at 900°C for one hour and air cooled, (2) austenitized at 845°C for one hour and oil quenched, and (3) tempered at 200°C for one hour and air cooled. This was done to obtain a relatively brittle material with a yield stress in uniaxial tension of $\sigma_0 \approx 1700$ MPa. Subsequently, one of the faces and all the edges of each specimen were ground to ensure flatness and perpendicularity. The ground surface was then lapped and polished to obtain an optically flat, highly reflective surface. The specimens had a nominal length of $2l =$ 0.30 m, width $b = 0.127$ m, and thickness $h = 0.01$ m with an initial notch length $a = 0.038$ m and root radius of 0.00015 m.

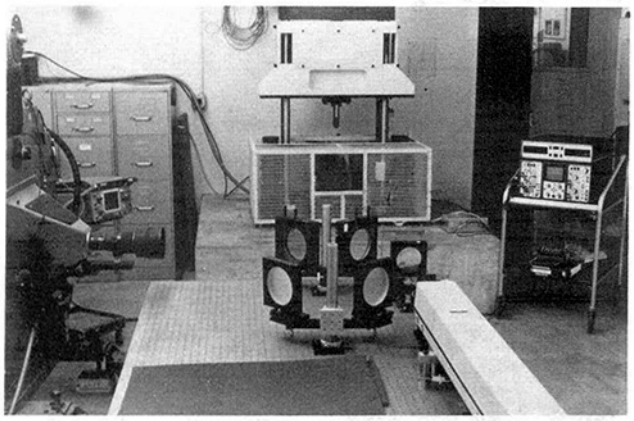


Fig. 4 Picture of experimental set-up

Loading Device. The specimens were loaded in a three-point bending configuration. The loading device used to dynamically load the specimens was the Dynatup 8100A drop-weight tower which works as follows: A drop weight whose mass and impact velocity can be varied to provide a range of impact energies, slides on two rails and can be hoisted to the desired height to obtain impact velocities of up to 10m/s. A pneumatic valve releases the weight from the chain in free fall. The hammer mounted at the bottom of the falling drop weight impacts the specimen which then gets dynamically loaded to the point of crack initiation, whence the crack propagates dynamically through the specimen.

To monitor the time-dependent tractions acting on the specimen boundaries during the impact event, both the hammer (the impact tup) and one of the supports (the support tup) are instrumented using strain gages and their readings are recorded on two channels of a Nicolet 4094 digital oscilloscope. These load records were used in finite element simulations of the experiments (as reported in Part II).

High-Speed Photography. A high-speed camera set-up (Fig. 4) was used to obtain a sequence of dynamic bifocal caustics. A Spectra Physics (model 166) argon-ion pulse laser was used as the light source. The laser beam was expanded and collimated to obtain a beam 100 mm in diameter. A reference grid was placed in the path of the collimated beam so that it lay in the object plane of the camera along one of the two bifocal paths. This grid provided the picture magnification. The laser beam that reflected off the polished specimen surface was split into two paths at the first 6-in. beam splitter of the bifocal set-up which consisted of an additional 6-in. beam splitter and two 6-in. first surface mirrors to provide a direct and a detoured path as described earlier. Both the beams were then collected into a rotating mirror, high-speed camera through a series of lenses. To obtain discrete pictures, the pulsing circuit of the laser was nominally set to give 25 ns exposure every 5 μs for a total of 1 ms from the time of an input trigger. Since the camera mirror was set to rotate at 800 rps, the above settings ensured that picture blur was negligible and also avoided overlap of frames while providing an adequate time window for filming the entire dynamic loading and crack propagation event. A mechanical interrupt-type photodiode switch mounted on the fixed tower of the drop-weight machine provided the trigger for the Nicolet oscilloscope, as well as for initiation of laser pulsing. This switch was positioned such that the falling drop weight itself would activate it at the moment of impact of the hammer with the specimen.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

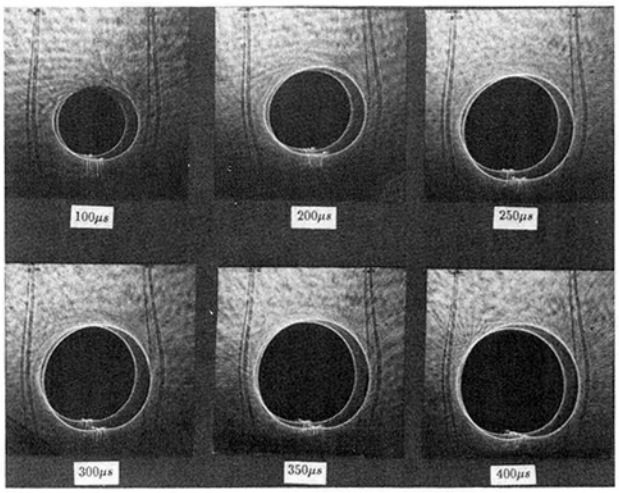


Fig. 5 Selection of bifocal caustics for a dynamically loaded stationary crack

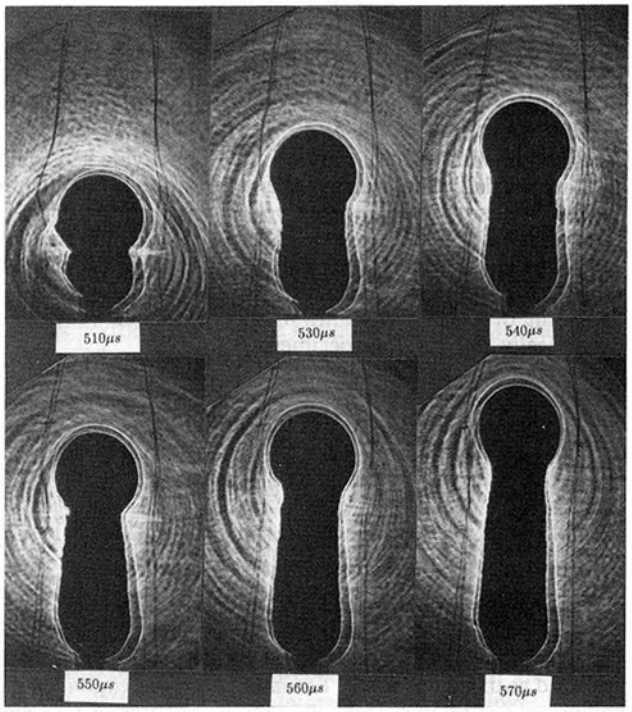


Fig. 6 Selection of bifocal caustics for a dynamically propagating crack

5 Results

Figure 5 shows a representative sequence of bifocal caustics for the case of the dynamically loaded stationary crack at various times from the time of impact. The superposition of the two images from the bifocal set-up was here so arranged as to have the pairs of caustics (the two bright lines around the crack-tip) one slightly shifted with respect to the other. Since a collimated laser beam was used as the light source, the smaller caustic corresponds to the smaller object plane distance (z_{02}) and the larger one to z_{01}. A selection of high-speed photographs for the case of the dynamically propagating crack is shown in Fig. 6. Here, the pairs of bifocal caustics are one inside the other, with the inner caustic corresponding to the smaller object plane distance (z_{02}). It is perhaps worth reiterating that the two bifocal images could be made to appear on the film track of the camera superposed, partially overlapping, or completely separated by simple translation of one of the mirrors of the bifocal set-up.

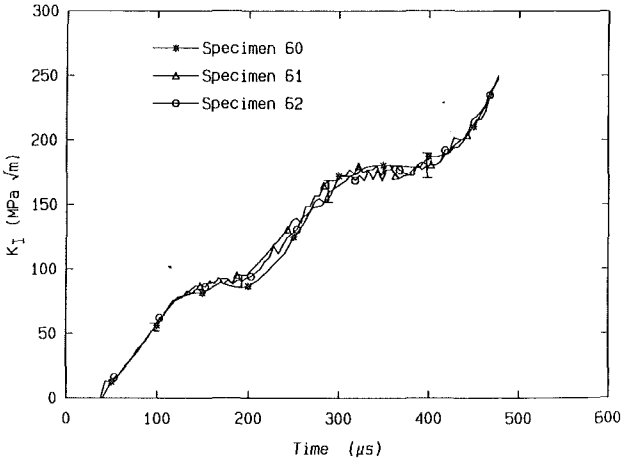


Fig. 7(a) Results for the loading phase from three identical tests

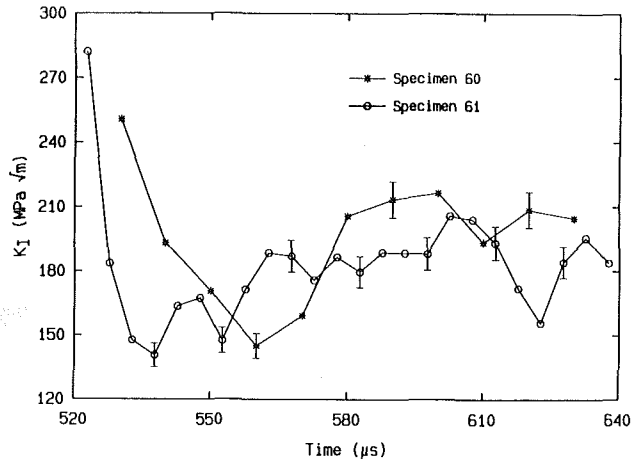


Fig. 7(b) Results for the propagating phase from two nominally identical tests

Measurement Error. To first get an idea of the repeatability and measurement errors involved in the method of caustics, a few experiments were conducted without using the bifocal technique to obtain caustics from a single z_0. Care was taken in these tests to have the specimen geometry, the loading and the object plane distance (z_0) identical. The results from the loading phase of three different tests are shown in Fig. 7(*a*). Here, as in the other figures, the time of impact is taken as zero time. A z_0 of 2.5 m was used for all these tests and the dynamic stress intensity factor was computed from the caustic diameters through equation (5). Note that the experimental scatter in the apparent stress intensity factor from three separate tests is seen to be about ± 5 percent. This is of the same order as the measurement error expected in a single test as obtained through an error analysis of equation (5). It should be emphasized that this is the error estimate *assuming* that equation (10) is valid. Thus, it should be recognized that differences in the stress intensity values from *bifocal caustics* that are less than the above value could arguably be attributed to measurement errors, while larger differences would be an indication of a more fundamental *malaise*—namely the possible violation of the assumptions underlying equation (10).

The situation for the dynamically propagating crack, however, is different. Figure 7(*b*) shows the measured dynamic stress intensity factor history during crack propagation from two ostensibly identical specimens. As was noted previously, the results for the loading phase of these tests were virtually identical to within experimental error. The crack initiation times (with reference to the time of impact), however, do not coincide. Thus, the rather large variations between the results of the two tests during the crack propagation phase

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

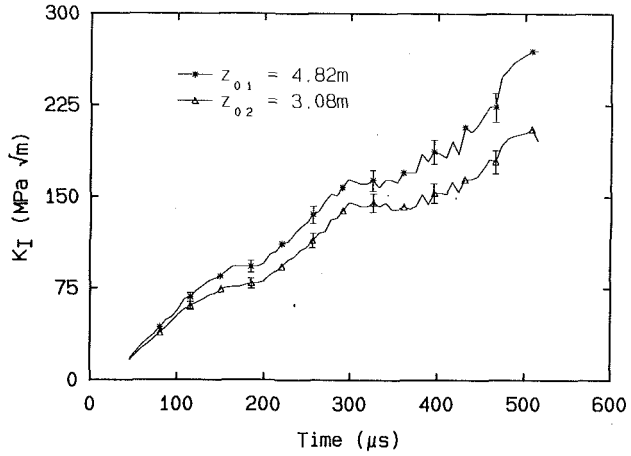


Fig. 8(a) Experimental results for the loading phase; specimen ($\alpha - 3$)

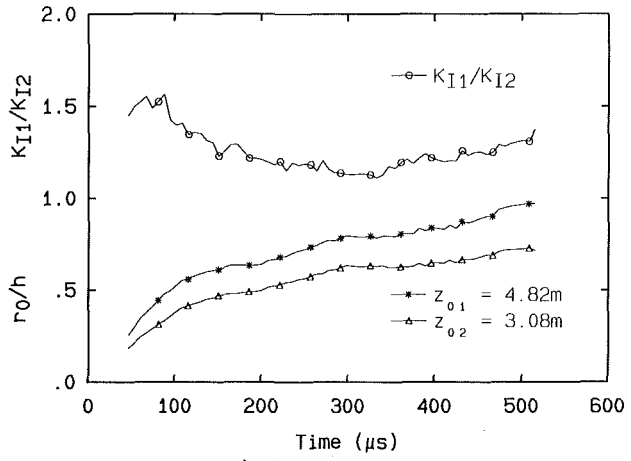


Fig. 8(b) Plot of K_{I1}^d/K_{I2}^d versus time for specimen ($\alpha - 3$)

arise not so much from experimental error as from the fact that the initiation times and, therefore, subsequent crack propagation histories are different. Unless the initial notch-tip bluntness is controlled to a degree of precision currently technically not feasible for metals, these differences are expected to be unavoidable. This, of course, is the primary reason for resorting to the bifocal technique which circumvents the necessity of comparing results from different tests.

Dynamically Loaded Stationary Crack. Turning now to the bifocal experiments, Fig. 8(*a*) shows the results for the uninitiated dynamically loaded crack for specimen ($\alpha - 3$). The two apparent stress intensity factors obtained from the diameters of the bifocal caustics pairs are shown as functions of time from impact. The object plane distances used in this experiment were $z_{01} = 4.82$m and $z_{02} = 3.08$m. The region of possible measurement uncertainty is indicated as vertical error bars. Figure 8(*b*) is an alternate representation of the experimental results for specimen ($\alpha - 3$). Here, the ratio K_{I1}^d/K_{I2}^d (see equation (12)) is plotted as a function of time from impact. As discussed previously, the experimental uncertainty in this ratio is less than in the individual values. For reference, the initial curve radii (as computed through equation (9)) corresponding to the measured stress intensity factor values for the two z_0's are also shown. It is seen that the apparent measured stress intensity factor is not quite independent of the region from which the measurement is made. Indeed, differences of up to 60 percent are observed between the measured stress intensity factors obtained from initial curves whose radii vary by less than 20 percent of the plate thickness. Unlike the results of Rosakis and Ravi-Chandar (1986) for the static case, these differences persist even for $r_0/h \geq 0.5$.

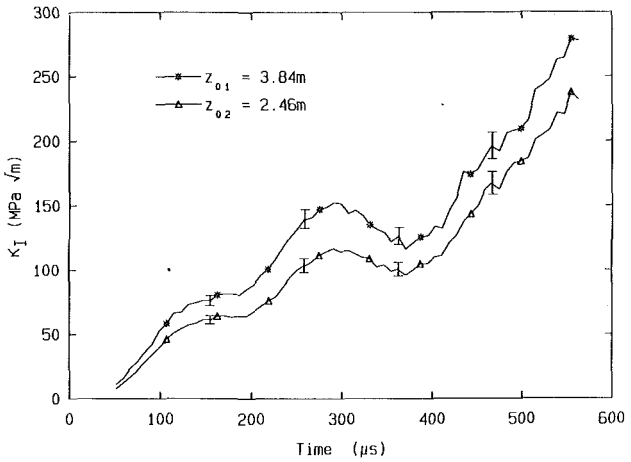


Fig. 9 Experimental results for the loading phase; specimen (v3s)

Essentially similar results were found in the other experiments (Fig. 9). Results from experiments on different specimens with various choices of z_0's all indicated (see Krishnaswamy (1988)) that the measured dynamic stress intensity factor increases quite substantially with increasing initial curve radius for the case of the dynamically loaded stationary crack. This has implications with regard to the extraction of dynamic initiation fracture toughness values.

Dynamically Propagating Crack. The results for the crack propagation phase are shown in Figs. 10–11. These are again time history plots of the two measured dynamic stress intensity factors from the bifocal caustics pairs. The larger initial curve is again seen to give a larger apparent stress intensity factor. Differences of up to 30 percent in the measured values are seen in these experiments. Note that in virtually all these tests, the initial curve radii are almost always greater than one-half the specimen thickness during the crack propagation phase. These differences, therefore, do not disappear for $r_0/h \geq 0.5$ unlike for the static case. Results for the propagating case with an even wider range of z_0's unfortunately could not be obtained because of the limitations of the experimental set-up.

With the results of both the dynamically loaded stationary crack and the dynamically propagating crack tests at hand, the following statement can be made about the dynamically loaded, three-point bend specimen: *the apparent dynamic stress intensity factor, as measured by the method of caustics, increases with increasing initial curve radius*. This means that the assumption that the K_I^d-field is dominant over this range of radius is not quite correct.

6 Conclusion

The experimental results for both stationary and propagating cracks lead to the conclusion that the near-tip dynamic out-of-plane displacement field seems to differ substantially from the asymptotic K_I^d-field. Indeed, these differences are such as to lead to large scatter in the measured apparent stress intensity factor above and beyond that caused by measurement errors. As will be discussed in detail in Part II, this scatter is of the same order as some of the observed variations cited in the literature as evidence of possible parametric dependence of the dynamic fracture toughness. In view of the results of the current experiments, one must wonder whether the observed variations in the literature might not also be due to violation of K_I^d-dominance presumed in the interpretation of the experimental data.

Some of the experiments to which the results of this work might have a bearing are briefly discussed here. The observed specimen dependence (Kalthoff; 1983) and acceleration

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

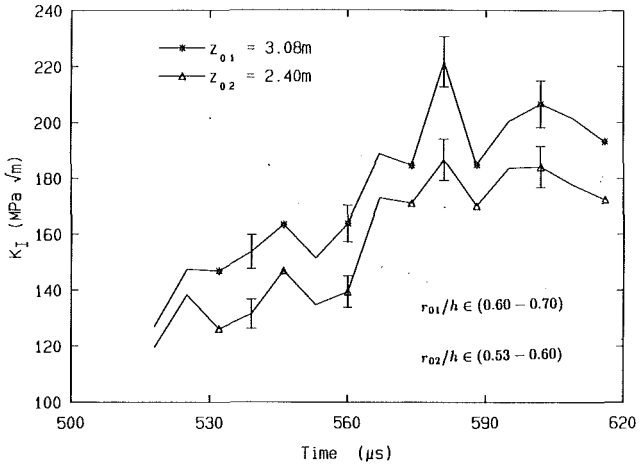


Fig. 10 Experimental results for crack propagation phase; specimen ($\alpha - 4$)

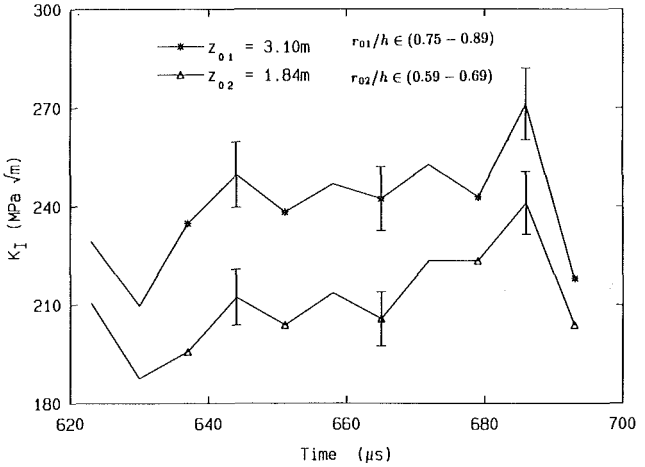


Fig. 11 Experimental results for crack propagation phase; specimen ($\alpha - 5$)

dependence (Takahashi and Arakawa; 1988) of the K_{ID}—$\dot{a}$ relation could possibly be manifestations of lack of K_I^d-dominance. This could also be the source of discrepancy in the results of Nigam and Shukla (1988), which show that the dynamic fracture toughness values obtained by the method of photoelasticity are 30–50 percent higher than those obtained through the method of caustics. These results, as well as others pertaining to the issue of uniqueness of the K_{ID}—$\dot{a}$ relation, will be discussed in greater detail in Part II after possible reasons for lack of K_I^d-dominance in the three-point bend specimens studied here have been identified.

Acknowledgments

The authors would like to thank Prof. G. Ravichandran for many helpful comments during the course of this study. We are grateful to Prof. R. J. Clifton for providing us with AISI 4340VAR steel. The contribution of Mr. Richard Pfaff toward upgrading of the high-speed camera at GALCIT is acknowledged. Support of ONR through contract N00014-85-K-0599 and NSF through PYI Grant MSM-84-51204 to the second author is gratefully acknowledged.

References

Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North Holland Publishing Company.

Baker, B. R., 1962, "Dynamic Stresses Created by a Moving Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, pp. 449–458.

Beinert, J., and Kalthoff, J. F., 1981, "Experimental Determination of Dynamic Stress-Intensity Factors by Shadow Patterns," *Mechanics of Fracture*, Vol. 7, G. Sih, ed., Sijthoff and Noordhoff, pp. 281–330.

Bilek, Z., 1980, "Some Comments on Dynamic Crack Propagation in a High Strength Steel," *Crack Arrest Methodology and Applications*, ASTM-STP 711, American Society for Testing and Materials, pp. 240–247.

Broberg, K. B., 1960, "The Propagation of a Brittle Crack," *Arkiv für Fysik*, Vol. 18, pp. 159–192.

Craggs, J. W., 1960, "On the Propagation of a Crack in an Elastic-Brittle Material," *Journal of Mechanics and Physics of Solids*, Vol. 8, pp. 66–75.

Dahlberg, L., Nilsson, F., and Brickstad, B., 1980, "Influence of Specimen Geometry on Crack Propagation and Arrest Toughness," *Crack Arrest Methodology and Applications*, G. T. Hahn et al., eds., ASTM-STP 711, pp. 89–108.

Dally, J. W., Fourney, W. L., and Irwin, G. R., 1985, "On the Uniqueness of K_{ID}—$\dot{a}$ Relation," *International Journal of Fracture*, Vol. 27, pp. 159–168.

Eshelby, J. D., 1969, "The Elastic Field of a Crack Extending Non-Uniformly under General Anti-Plane Loading," *Journal of Mechanics and Physics of Solids*, Vol. 17.

Freund, L. B., 1972, "Crack Propagation of a Crack in an Elastic Solid Subjected to General Loading-I. Constant Rate of Extension," *Journal of Mechanics and Physics of Solids*, Vol. 20, pp. 129–140.

Freund, L. B., and Clifton, R. J., 1974, "On the Uniqueness of Plane Elastodynamic Solutions for Running Cracks," *Journal of Elasticity*, Vol. 4, No. 4, pp. 293–299.

Freund, L. B., and Hutchinson, J. W., 1985, "High Strain-Rate Crack Growth in Rate-Dependent Plastic Solids," *Journal of Mechanics and Physics of Solids*, Vol. 33, pp. 169–191.

Kalthoff, J. F., 1983, "On Some Current Problems in Experimental Fracture Mechanics," *Workshop on Dynamic Fracture*, W. G. Knauss et al., eds., California Institute of Technology, pp. 11–35.

Kalthoff, J. F., 1985, "On the Measurement of Dynamic Fracture Toughness—A Review of Recent Work," *International Journal of Fracture*, Vol. 27, pp. 277–298.

Knauss, W. G., and Ravi-Chandar, K., 1985, "Some Basic Problems in Stress Wave Dominated Fracture," *International Journal of Fracture*, Vol. 27, pp. 127–143.

Kobayashi, A. S., and Mall, S., 1978, "Dynamic Fracture Toughness of Homalite 100," *Experimental Mechanics*, Vol. 18, pp. 11–18.

Kobayashi, T., and Dally, J. W., 1977, "Relation Between Crack Velocity and the Stress Intensity Factor in Birefringent Polymers," *Fast Fracture and Crack Arrest*, ASTM-STP 627, G. T. Hahn et al., eds., pp. 257–273.

Krishnaswamy, S., 1988, "On the Domain of Dominance of the Asymptotic Elastodynamic Crack-Tip Fields," doctoral dissertation, California Institute of Technology.

Lam, P. S., and Freund, L. B., 1985, "Analysis of Dynamic Crack Growth of a Tensile Crack in an Elastic-Plastic Material," *Journal of Mechanics and Physics of Solids*, Vol. 33, pp. 153–167.

Ma, C. C., and Freund, L. B., 1986, "The Extent of the Stress Intensity Factor During Crack Growth under Dynamic Loading Conditions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 303–310.

Manogg, P., 1964, "Anwendungen der Schattenoptik zur Untersuchung des Zerreissvorgags von Platten," doctoral dissertation, Freiburg, W. Germany.

Nakamura, T., and Parks, D. M., 1987, "Three-Dimensional Field Near the Crack Front of a Thin Elastic Plate," presented at the 24th Annual Meeting of the Society of Engineering Science, Sept. 21–23, 1987, Salt Lake City, Utah.

Nigam, H., and Shukla, A., 1988, "Comparison of the Techniques of Transmitted Caustics and Photoelasticity as Applied to Fracture," *Experimental Mechanics*, Vol. 28, No. 2, pp. 123–133.

Parsons, I. D., Hall, J. F., and Rosakis, A. J., 1986, "A Finite Element Investigation of the Elastostatic State Near a Three-Dimensional Edge Crack," GALCIT Report SM 86-29, California Institute of Technology.

Ravi-Chandar, K., and Knauss, W. G., 1984, *International Journal of Fracture*, Vol. 25, pp. 247–262.

Ravi-Chandar, K., and Knauss, W. G., 1987, "On the Characterization of the Transient Stress Field Near the Tip of a Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 72–78.

Rosakis, A. J., Duffy, J., and Freund, L. B., 1984, "The Determination of Dynamic Fracture Toughness of AISI 4340 Steel by the Shadow Spot Method," *Journal of the Mechanics and Physics of Solids*, Vol. 32, pp. 443–460.

Rosakis, A. J., and Zehnder, A. T., 1985, "On the Method of Caustics: An Exact Analysis Based on Geometrical Optics," *Journal of Elasticity*, Vol. 15, No. 4, pp. 347–368.

Rosakis, A. J., and Ravi-Chandar, K., 1986, "On Crack-Tip Stress State: An Experimental Evaluation of Three-Dimensional Effects," *International Journal of Solids and Structures*, Vol. 22, No. 2, pp. 121–134.

Smith, R. H., and Freund, L. B., 1980, "Three-Dimensional Finite Element Analysis of Steady Elastodynamic Crack Growth in a Plate," *Analytical, Numerical and Experimental Aspects of Three-Dimensional Fracture Processes*, A. J. Rosakis et al., eds., ASME-AMD-Vol. 91, pp. 185–196.

Takahashi, K., and Arakawa, K., 1987, "Dependence of Crack Initiation on the Dynamic Stress-Intensity Factor in Polymers," *Experimental Mechanics*, Vol. 27, No. 2, pp. 195–199.

Yang, W., and Freund, L. B., 1985, "Transverse Shear Effects for Through Cracks in an Elastic Plate," *International Journal of Solids and Structures*, Vol. 21, No. 9, pp. 977–994.

Yoffe, E. H., 1951, "The Moving Griffith Crack," *Philosophical Magazine*, Vol. 42, pp. 739–750.

Zehnder, A. T., and Rosakis, A. J., 1989, "Dynamic Fracture Initiation and Propagation in 4340 Steel Under Impact Loading," *International Journal of Fracture*, (to appear).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Sridhav Krishnaswamy
Research Fellow.

Ares J. Rosakis
Assoc. Professor,
Assoc. Mem. ASME

Graduate Aeronautical Laboratories,
California Institute of Technology,
Pasadena, CA 91125

G. Ravichandran
Asst. Professor,
University of California,
San Diego, CA
Assoc. Mem. ASME

On the Extent of Dominance of Asymptotic Elastodynamic Crack-Tip Fields: Part II—Numerical Investigation of Three-Dimensional and Transient Effects

In Part I of this paper, the question of the extent of dominance of the mode I asymptotic elastodynamic crack-tip field (the K_I^d-field) was studied experimentally. Here, the results of two and three-dimensional elastodynamic finite element simulations of the drop-weight experiments are reported. The load records as obtained from the impact hammer and supports of the drop-weight loading device were used as boundary tractions in the numerical simulations. For the laboratory specimen studied, the results of the simulations indicate that the asymptotic elastodynamic field is not an adequate description of the actual fields prevailing over any sizeable region around the crack tip. This confirms the experimental results of Part I which showed that three-dimensional and transient effects necessarily have to be taken into account for valid interpretation of experimental results.

1 Introduction

It is increasingly being recognized that the stress and displacement fields near the tip of a crack are essentially three-dimensional in nature. Exact analytical solutions for the three-dimensional crack problem, however, do not exist in general. Thus, finite elements have been used extensively to gain an understanding of the three-dimensional structure of the crack-tip fields. For example, Levy, Marcal, and Rice (1971), Parsons, Hall, and Rosakis (1986), and Nakamura and Parks (1987) have investigated three-dimensional crack problems under elastostatic conditions. In addition, Narasimhan and Rosakis (1988) and Nakamura and Parks (1988) have recently studied the three-dimensional elastoplastic state prevailing near the tip of a static crack.

The finite element method has also been used to study three-dimensional dynamic fracture problems. For the most part, however, special singular elements have been used to model the crack tip. The purpose of employing singular elements has been to *bias* the near-tip field to have a particular (asymptotically known) structure in order to extract the magnitude of crack-tip singularities (the stress intensity factor, J-integral, etc). For instance, Nakamura, Shih, and Freund (1986) have analyzed the three-point bend configuration under idealized dynamic loading conditions. Their primary objective was to identify conditions under which experimentally obtained quantities could be used to directly infer the initiation value of the J-integral under elastoplastic conditions.

Smith and Freund (1988) were the first to investigate in detail the nature of the near-tip three-dimensional elastic fields for a steady-state, dynamically propagating crack under Mode I conditions. They, however, *imposed* the plane-stress dynamic stress intensity factor field (the K_I^d-field) as the far-field boundary condition. As pointed out in Part I of this paper, one of the most important pieces of information that the experimentalist would like to have is the extent to which the K_I^d-field models the near-tip stresses and displacements. Most of the experimental results reported in the literature rely on the assumption of K_I^d-dominance around the crack tip. The adequacy of such an assumption therefore needs to be studied.

In Part I, the issue of K_I^d-dominance was studied through a series of drop-weight experiments in conjunction with the method of *bifocal caustics*. The results of the experiments indicated that the assumption of K_I^d-dominance could lead to substantial errors in the measured results. In this paper, the finite element method is used to investigate this issue further. No attempt is made to bias the near-tip singularity; nor is the far-field boundary condition constrained to be given by the plane-stress asymptotic field.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 31, 1989; final revision, Mar. 14, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A series of two and three-dimensional finite element simulations[1] of the drop-weight experiments (reported in Part I) were done using the experimentally recorded impact and support-point load histories as boundary tractions for the simulations. The simulations were undertaken in an effort to qualitatively capture the essential features of the experimental results. Only the case of the dynamically loaded stationary crack was considered for the simulations. It was felt that attempting to simulate the case of a propagating crack would be premature at this point, especially since it entails a high degree of uncertainty in terms of the nodal-release procedure that would have to be used. The primary interest here is in trying to identify the role of specimen geometry, dynamic loading, and the three-dimensional structure of the crack-tip region (especially under transient conditions) insofar as these have a bearing on the K_I^d-dominance assumption on which the experimental methods rest.

Three issues are addressed. First, where relevant, direct comparisons of the numerical results with the corresponding asymptotic field are made to ascertain the adequacy of the K_I^d-field as a characterizer of the near-tip continuum fields. All field quantities presented are normalized in such a manner as to highlight salient points. Thus, full-field two-dimensional results are compared with the asymptotic field and three-dimensional results are normalized by the appropriate two-dimensional or asymptotic values. Secondly, virtual energy release rate integrals evaluated numerically are used to extract stress intensity factor values in order to compare with the experimentally measured values. Finally, the implications of the finite element results are studied with regard to the experimental method of caustics in reflection. In an attempt to qualitatively recover the results of the drop-weight experiments, an exact analog of the procedure used in the experiments is attempted in the following manner. As explained in Part I, the out-of-plane surface displacement fields obtained from the analyses are used to numerically generate (synthetic) caustic patterns for a range of initial curves. That is, the numerically obtained surface out-of-plane displacement field u_3 $(x_1, x_2, h/2)$ is subjected to the optical mapping

$$\mathbf{X} = \mathbf{x} + 2z_0 \nabla u_3(x_1, x_2, h/2) \tag{1}$$

for a set of values for parameter z_0 (the object plane distance). Here, $\mathbf{x}$ are points on the specimen that are optically mapped onto points $\mathbf{X}$ on the object plane, and h is the specimen thickness. These caustics are then interpreted exactly as in the experiments; i.e., the caustic diameters are related to the stress intensity factor under the assumption that the underlying out-of-plane displacement field is K_I^d-dominant. Thus, the transverse diameters of the synthetic caustics are used to extract the apparent stress intensity factor values through the relation

$$K_I^d = \frac{ED^{5/2}}{z_0 \nu h} f(\dot{a}, \nu), \tag{2}$$

where D is the transverse diameter of the caustic, ν is Poisson's ratio and E is the Young's modulus of the material, h is the specimen thickness, $\dot{a}$ is the crack velocity, and f is a known function of crack velocity and material properties (see Part I for details). Again, as in the experiments, if the displacement field is not actually K_I^d-dominant, this fact would be reflected as an apparent (erroneous) dependence of the stress intensity factor on the radial distance from the crack front.

It must be pointed out that in the ensuing discussion of two-dimensional results, all in-plane lengths are normalized by the (actual) specimen thickness, h. For a two-dimensional problem, however, the plate thickness is not a relevant length scale. This is done here purely for ease of reference with subsequent three-dimensional results.

2 Two-Dimensional Elastodynamic Simulations

The simulations of the dynamic experiments were first attempted under the simplifying assumption that the specimens could be considered to be essentially under plane-stress conditions. Before delving into the simulations of the experiments, one issue needs to be addressed. Since no special singular elements were used in the finite element analyses, it is essential that the discretization used must be such as to capture the expected singular crack-tip fields adequately. To this effect, preliminary two-dimensional elastostatic analyses of the three-point bend specimen were performed. Based on the results of these, the mesh discretization shown in Fig. 1 consisting of 396 isoparametric linear quadrilateral elements (425 nodes) was found to be adequate. This mesh has a focused region around the crack-tip of about one (actual) specimen thickness which is divided into 18 sectors and 10 concentric rings of elements. The crack-tip elements are four-noded quadrilaterals collapsed into triangles. The mesh discretization was found to adequately capture the square-root singular field near the crack tip for the elastostatic case and was thus deemed suitable for the dynamic problem as well.

For the two-dimensional elastodynamic simulations, the loads as obtained from the experimental tup records (Fig. 2) were applied as the boundary conditions. That is, the impact-tup load history was applied to the nodes corresponding to the impact tup, and the support-tup load history was applied to the associated nodes as shown in Fig. 1. From symmetry conditions, the uncracked ligament was constrained to move only along the x_1-direction. The rest of the boundary was left free

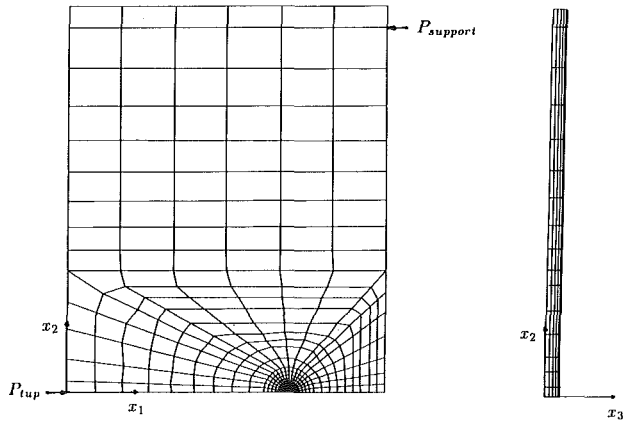


Fig. 1 Mesh geometry

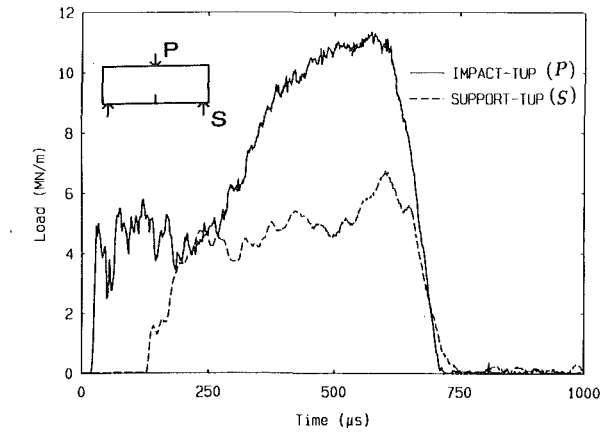


Fig. 2 Impact and support load records; specimen (α – 4)

[1]All the numerical computations were done using a much modified version of the finite element analysis program FEAP whose ancestry traces back to Taylor (1977).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

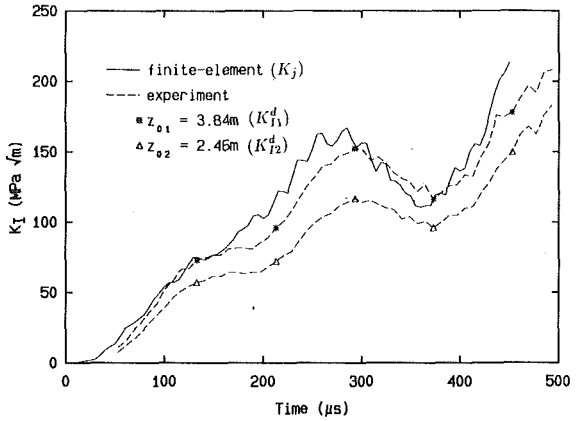


Fig. 3(*a*) Comparison of experimentally obtained dynamic stress intensity factor history with that computed from a two-dimensional simulation; specimen (*v* 3*s*)

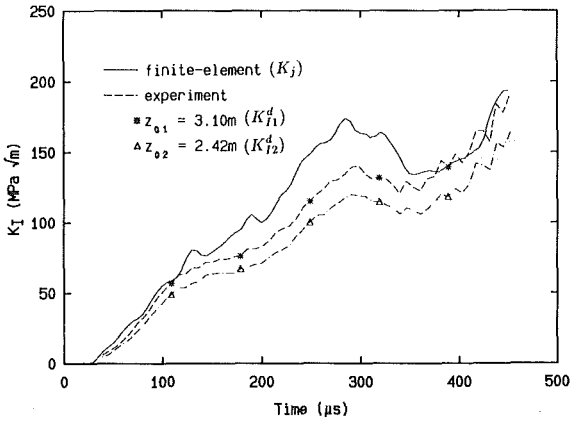


Fig. 3(*b*) Comparison of experimentally obtained dynamic stress intensity factor history with that computed from a two-dimensional simulation; specimen (α – 4)

of traction. An implicit Newmark predictor-corrector time integration scheme (see Hughes and Belytschko, 1983) was used for its virtue of unconditional stability which would allow for relatively large time steps and the attendant loss of high-frequency information was deemed acceptable since it is not the intent here to monitor discrete stress waves in the body. The time steps chosen for the numerical integration were thus based only on accuracy considerations.

The (virtual) energy release rate for a dynamically loaded stationary crack is given by (see Nakamura, Shih, and Freund, 1986)

$$J=\lim_{\Gamma\to 0}\int_{\Gamma}\left((U+T)n_1-\sigma_{ij}n_j\frac{\partial u_i}{\partial x_1}\right)d\Gamma, \tag{3}$$

where U is the strain energy density, T is the kinetic energy density, σ_{ij} is the stress tensor, $\mathbf{u}$ is the displacement vector, and $\mathbf{n}$ is the unit outward normal to the contour of integration Γ. Here, $\Gamma \to 0$ symbolically indicates that the integration contour must be shrunk on to the crack tip. In the simulations, the time history of this integral was computed using the equivalent domain integral form described in Nakamura, Shih, and Freund (1986). The dynamic stress intensity factor was then computed through the relation between K^d_I and J in plane stress,

$$K^d_I = \sqrt{EJ}, \tag{4}$$

where E is the Young's modulus of the material. Figures (3*a*), (3*b*) show the experimentally obtained dynamic stress intensity factor history in comparison with that from the numerical simulations for specimens ($v3s$) and ($\alpha-4$). Here, K^d_{I1} and K^d_{I2}

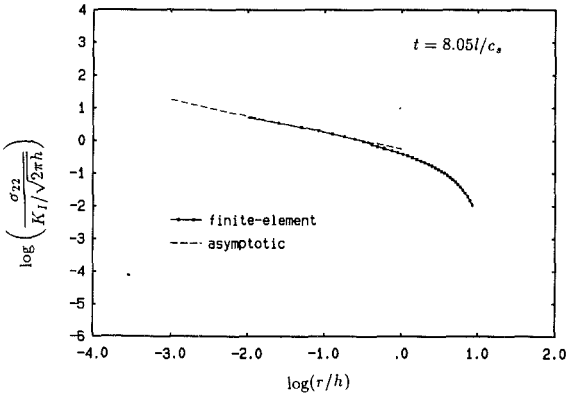


Fig. 4 Plot of log ($\sigma_{22}\sqrt{2\pi h}/KK^{\dagger}(t)$) versus log ($r/h$) for the two-dimensional elastodynamic case

refer to the experimentally measured values corresponding to the two object plane distances z_{01} and z_{02} (see Part I for details) and K_j is that computed from the dynamic simulations (through the J-integral). It is seen that in both cases K_j has the same general trend as K_{I1} and K_{I2} except that the experimental values are sometimes substantially lower, while at other times equal to or higher than the simulated values. This discrepancy is attributable to two sources. First, there are uncertainties associated with the simulations in terms of how accurately the tup records provide the boundary tractions actually experienced by the specimens. Secondly, and more importantly, there is the possibility that the experimental values might not have been obtained from a region of K^d_I-dominance. This is in fact foreshadowed by the discrepancy between the two experimental records themselves.

In the above, it has been implicitly assumed that the asymptotic K^d_I-field has validity for this (two-dimensional) geometry and dynamic loading condition. However, as pointed out in Part I, the existence of a stress intensity factor field around a dynamically loaded stationary crack in a finite geometry has not been universally established. It is thus necessary to check whether a square-root singular asymptotic field is appropriate here. To this end, a logarithmic plot of stress component σ_{22} along the $\theta = 5$ deg line versus the logarithm of radial distance from the crack tip is shown in Fig. 4 for one representative time in the simulation. Comparison with the corresponding curves for the asymptotic field (with the values of the stress intensity factor obtained from the J-integral) indicates that a square-root singular field is indeed asymptotically descriptive of the near-tip continuum structure for this two-dimensional configuration at least for the times shown.

As a measure of the extent of dominance of the asymptotic K^d_I-field, the angular variation of the stresses and displacements for a range of radial distance is shown in Figs. 5(*a*), 5(*b*) for one particular time. Also shown for comparison are the corresponding asymptotic values. Note that the normalization used here is such that the asymptotic values are given by a single curve for *any* radial distance. This normalization enables distinct features of the stress intensity factor field to be discerned in the near-tip full-field solution. The magnitudes, however, are seen to vary with increasing radial distance from the crack tip.

A quantity of fundamental interest for the method of caustics in reflection is the out-of-plane displacement field, u_3. The plane-stress condition that $\sigma_{33} = 0$ provides the following relation for the out-of-plane displacement field:

$$u_3 = -\frac{\nu h}{2E}(\sigma_{11}+\sigma_{22}), \tag{5}$$

where h is the actual specimen thickness. The angular variation of this is shown in Fig. 6 for two radial distances from the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

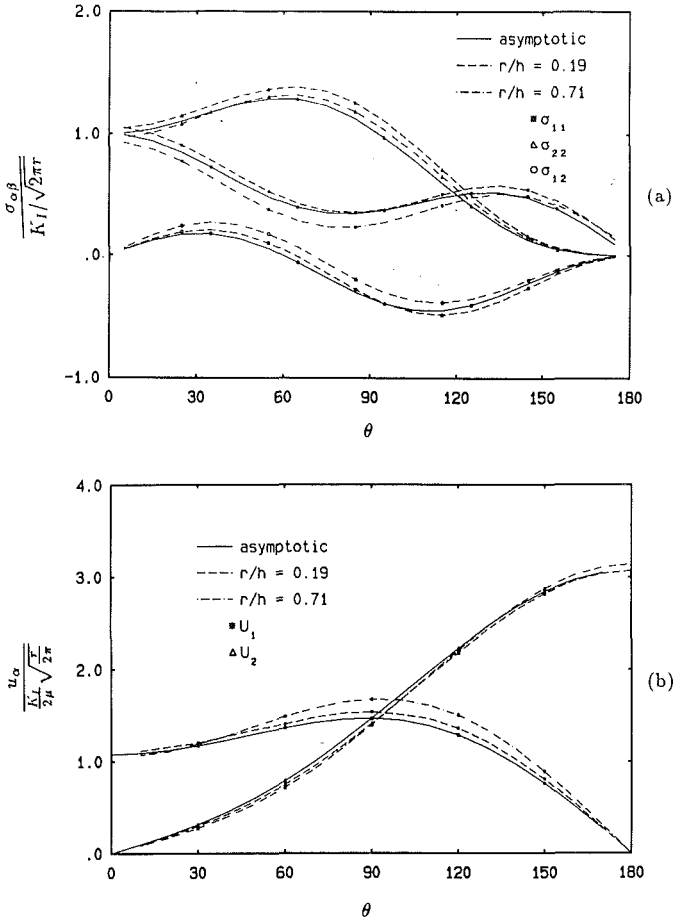


Fig. 5 Angular variation of near-tip (*a*) stresses and (*b*) in-plane displacements for the two-dimensional case in comparison with the corresponding asymptotic values

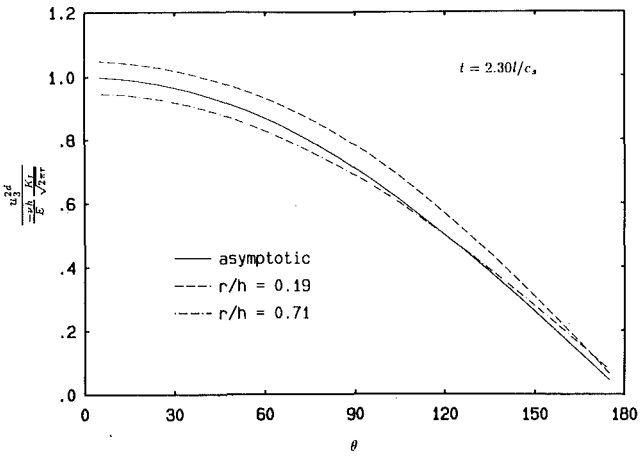


Fig. 6 Angular variation of computed out-of-plane displacements for the two-dimensional case in comparison with the corresponding asymptotic values

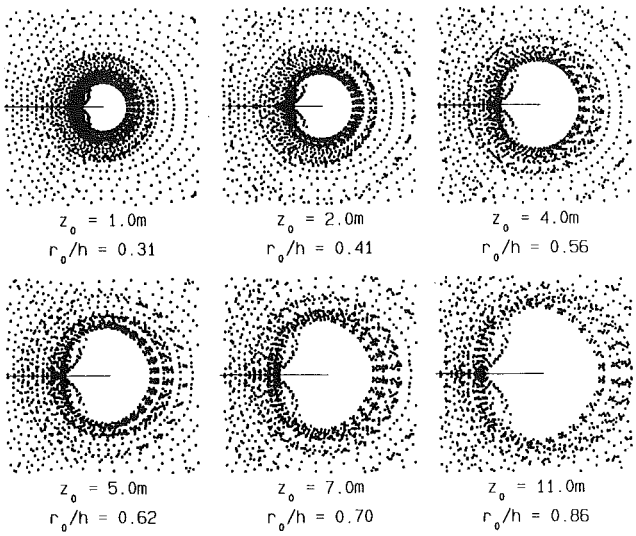


Fig. 7(*a*) Numerically generated caustics from the two-dimensional elastodynamic simulation

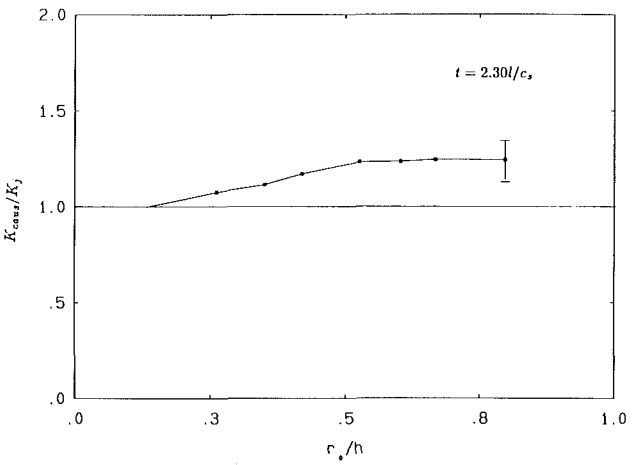


Fig. 7(*b*) Plot of K_{caus}/K_j versus r_0/h from the two-dimensional elastodynamic simulation

crack tip. Again, u_3 is normalized by the corresponding asymptotic quantity. It can be seen that the full-field quantity is in reasonable qualitative agreement with the asymptotic expression for $r/h \rightarrow 0$ with increasing deviation in magnitude as r/h increases. This was seen to be the case for other times in the simulation as well. Note, again, that normalization of two-dimensional results with respect to the (actual) specimen thickness is done purely for comparison purposes with subsequent three-dimensional results.

Finally, synthetic caustics were obtained from the out-of-plane displacement field for various times in the simulation. Since linear finite elements were used and, in plane-stress problems, the out-of-plane displacement field is obtained from the in-plane stress field, a smoothing scheme given in Hinton and Campbell (1974) was used to obtain the derivatives of the surface-displacement field required in the caustic mapping of the surface. A representative set of these caustic patterns is shown in Fig. 7(*a*) for one particular time in the simulation and for a set of $z_0 s$. The corresponding initial-curve radii computed on the basis of the caustic diameters (see Part I for details) range from r_0/h of 0.31 to 0.86. These caustics are seen to have the epicycloidal shape characteristic of those from the asymptotic solution. The ratio of the stress intensity factor computed from the diameters of these synthetic caustics (denoted K_{caus}) to that obtained from the J-integral (K_j) are plotted in Fig. 7(*b*) as a function of the initial curve radius of the numerical caustic. While the apparent measured stress intensity factor seems to increase as r_0/h increases, to within the accuracy warranted by the procedure used here, it appears that caustics should provide the stress intensity factor value (to within 20 percent) for initial curve radii in the range $r_0/h \leq 0.8$. Even though this error is substantial, it is clear that two-dimensional transient effects alone would not seem to entirely account for the much larger variation in the stress intensity factor observed in the experiments. Indeed, it is worth noting that the above

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

result would indicate that caustics should always overestimate K_j which is not necessarily the case in the experiments (see Figs. 3(*a*), 3(*b*)).

3 Three-Dimensional Elastodynamic Simulation

It appears that the substantial variation observed between the experimentally measured stress intensity factors from bifocal caustics (see Part I) cannot be explained purely in terms of (two-dimensional) dynamic effects affecting the caustic patterns. Thus, additional reasons must be sought in terms of (a) nonlinear effects or (b) three-dimensional effects under transient conditions. Visual evidence of the plastic deformation in the fractured specimens indicated that the initial curves for the experimental caustics were well outside the plastic zone whose *maximum extent* was seen to be confined to $r_p/h \leq 0.15$ (Krishnaswamy, 1988). Based on the estimates of Rosakis and Freund (1981), the experimental results of Zehnder and Rosakis (1988), as well as elastoplastic simulations of the current experiments, plasticity effects were found to be minimal. Thus, attention will now be directed toward studying the effect of three-dimensionality near the crack tip by means of a full-field three-dimensional elastodynamic simulation of the drop-weight experiments.

The mesh geometry used had an in-plane layout identical to that used for the two-dimensional cases (Fig. 1). This enables direct comparison of three-dimensional results with the corresponding plane-stress simulations and thereby helps identify the effect of three-dimensionality. Five layers of 8-noded brick elements through half the thickness leading to a total of 1960 elements (2550 nodes) were used to model one-quarter of the three-dimensional body. The crack-tip elements were collapsed to form triangular wedges. Recognizing that the largest through-thickness variations in field quantities occur near the free surface, the mesh was graded in the thickness direction such that the layer interfaces were at $x_3/h = 0$, 0.15, 0.275, 0.375, 0.45, and 0.5. The experimentally obtained tup load histories interpreted as uniform line loads through the thickness were applied as boundary conditions to the appropriate nodes. The uncracked ligament surface and the specimen midplane were constrained suitably as dictated by symmetry considerations. The rest of the boundary was left traction-free. Once again, an implicit Newmark predictor-corrector scheme was used in order to be consistent with the algorithm used for the two-dimensional simulation.

Following Nakamura, Shih, and Freund (1986), an average dynamic energy release rate integral appropriate for the three-dimensional case can be defined as:

$$J_{av} = \frac{1}{h} \lim_{S \to 0} \int_S \left((U+T) n_1 - \sigma_{ij} n_j \frac{\partial u_i}{\partial x_1} \right) dS, \qquad (6)$$

where S is now a tubular surface through the specimen and $S \to 0$ symbolically indicates that this surface is shrunk onto the crack front. In order to compare with experimental results, an "average" stress intensity factor is extracted from J_{av} through relation (4). The average energy release rate value obtained from the three-dimensional simulation is shown in Fig. 8 and is seen to be not much different from that computed in the plane-stress analysis. This suggests that the plane-stress approximation might be adequate if one is merely interested in integrated energy release rate type of quantities.

From the point of view of the method of caustics in reflection, the primary issue is the extent of deviation of the near-tip surface out-of-plane displacement field from the corresponding asymptotic plane-stress expression. This is shown in Fig. 9(*a*) for a set of radial lines along $\theta = 0$ deg, 60 deg, and 140 deg and for one typical time. The surface u_3-displacements as obtained from the three-dimensional solution are normalized by the corresponding asymptotic values. The two-dimensional counterpart along the radial line $\theta = 0$ deg is shown for comparison purposes. Two salient features can be identified from this figure. First, the *two-dimensional* full-field values seem to be in good agreement with the asymptotic for sufficiently small radial distances. The asymptotic expression becomes increasingly inaccurate for larger radial distances. Secondly, it is precisely in those regions where the two-dimensional and asymptotic fields are in close agreement that the three-dimensional structure deviates the most from the asymptotic. In Fig. 9(*b*), the angular variations of the surface u_3-displacements are shown for a range of r/h. Under the normalization used, the asymptotic u_3-displacement are given by $\cos(\theta/2)$—the solid line—for any radius. Again, it is noted that the asymptotic field does not always model the actual three-dimensional structure very well. Indeed, the deviation from the asymptotic is seen to be pronounced toward the crack tip. For larger r/h, the theta variation of u_3 seems to approach the asymptotic curve in form, though not in actual magnitude. Further, Fig. 9(*c*) clearly brings out the deviation of the actual three-dimensional field from the corresponding asymptotic values. Here, contours of constant free-surface u_3-displacements from the three-dimensional analysis are shown in the top half of the figure. The bottom half shows the *corresponding* contours as if the asymptotic K_I^d-field had prevailed. It appears that the actual three-dimensional field is not quite captured by the asymptotic expression for the surface out-of-plane displacements—toward the crack tip because of substantial three-dimensional effects and away from it because the asymptotic expression becomes increasingly insufficient.

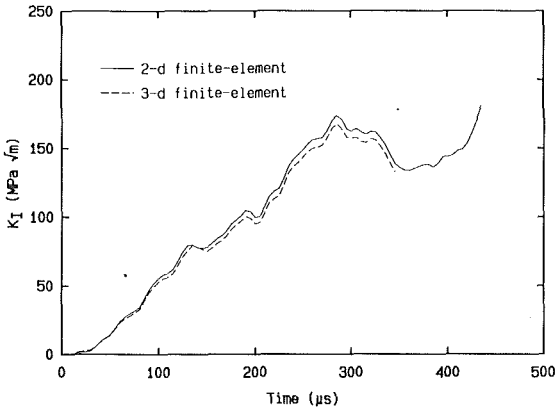


Fig. 8 Comparison of "average" dynamic stress intensity factor history as computed from two and three-dimensional simulations

The qualitative effects of the above features insofar as these affect the method of caustics can now be evaluated. As described previously, the three-dimensional surface out-of-plane displacements obtained numerically for various times in the simulation are mapped using equation (1) to obtain synthetic caustic patterns. A representative sequence of these caustic patterns for one particular time is shown in Fig. 10(*a*). These were obtained for the same instant of simulation and for the same set of z_0s as were used in the two-dimensional case. The role of three-dimensionality is thus clearly brought out. The graphic contrast that emerges between the caustic patterns obtained under two-dimensional assumptions and the actual three-dimensional conditions can be seen by comparing Figs. 7(*a*) and 10(*a*). In comparison to the corresponding two-dimensional results, it is noted that for small r_o/h, the shadow spots in Fig. 10(*a*) are less epicycloidal in shape and much smaller in size. This is not surprising considering the angular variations of u_3 which deviated markedly from the asymptotic value as $r/h \to 0$. Compare this with the results of the plane-stress analysis (see Fig. 7(*a*)) where the qualitative agreement with the asymptotic field appears to be much better over a wider range of r/h. If, now, one were to relate the caustic

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

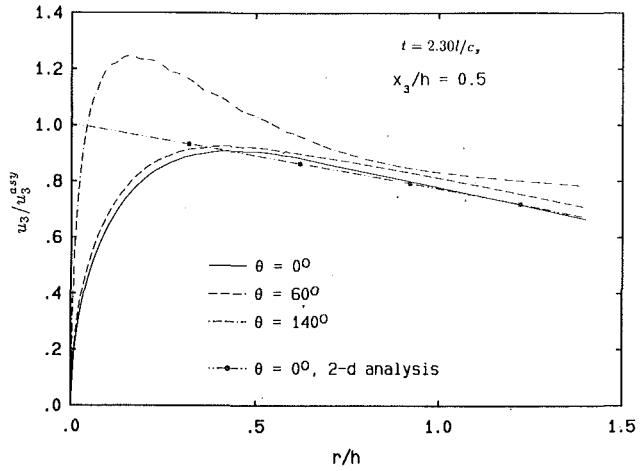


Fig. 9(*a*) Radial variation of the free surface u_3-displacements in comparison with the corresponding asymptotic values

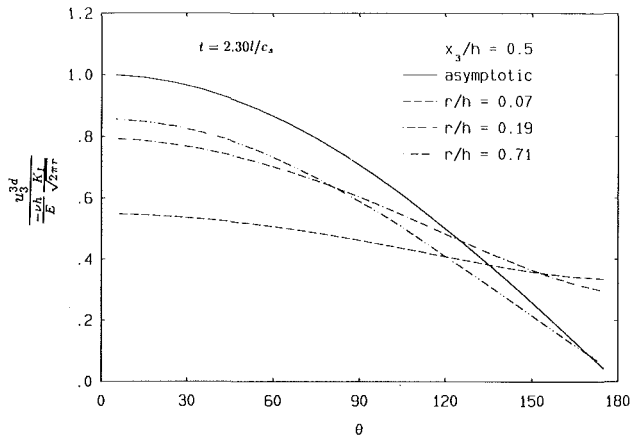


Fig. 9(*b*) Angular variation of the free surface u_3-displacements in comparison with the corresponding asymptotic values

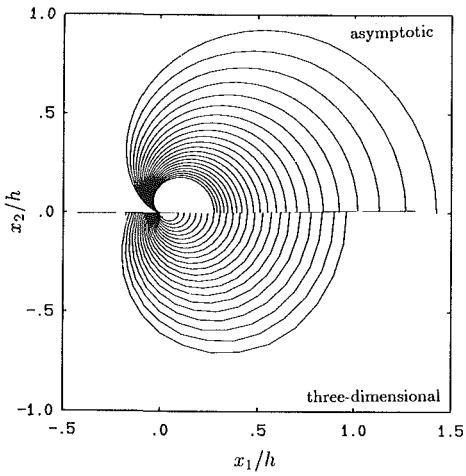


Fig. 9(*c*) Contours of constant free surface u_3-displacements in comparison with the corresponding asymptotic values

diameters to the stress intensity factor through equation (2), then the resulting value of K_{caus}/K_j (shown for two times) is seen to vary quite substantially with increasing initial curve radius as shown in Fig. 10(*b*). From the figure, it is seen that as $r_0/h \to 0$, the measured stress intensity factor value becomes substantially less than the value obtained from the domain integral. Further, for larger initial curve radii, it is possible for caustics to provide an overestimate of K_j. More importantly, the almost monotonically increasing K_{caus}/K_j versus r_0/h curve for the three-dimensional transient simulation case qualitatively captures essentially all the features observed experimentally. Thus, the apparent dynamic stress intensity factor, as measured by the method of caustics, would seem to increase with increasing initial-curve radii. Also, in view of the lack of a sizeable domain of dominance of the K_I^d-field, it would appear that the agreement between the measured dynamic stress intensity factor history and that computed through the *J*-integral cannot be expected to be any better than obtained in Figs. 3(*a*), 3(*b*).

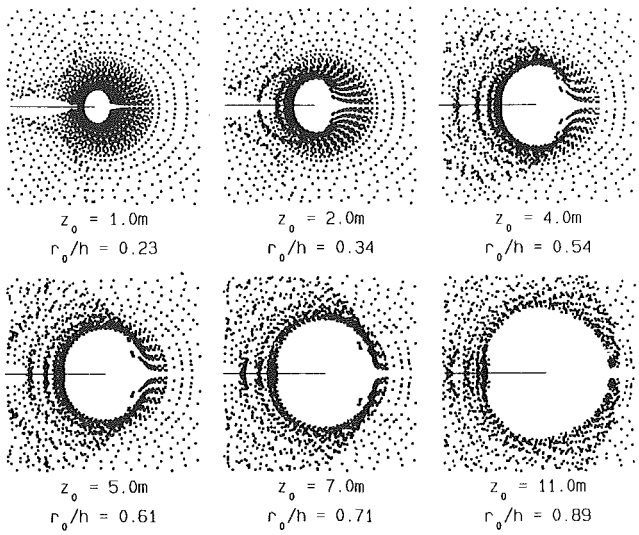


Fig. 10(*a*) Numerically generated caustics from the three-dimensional elastodynamic simulation

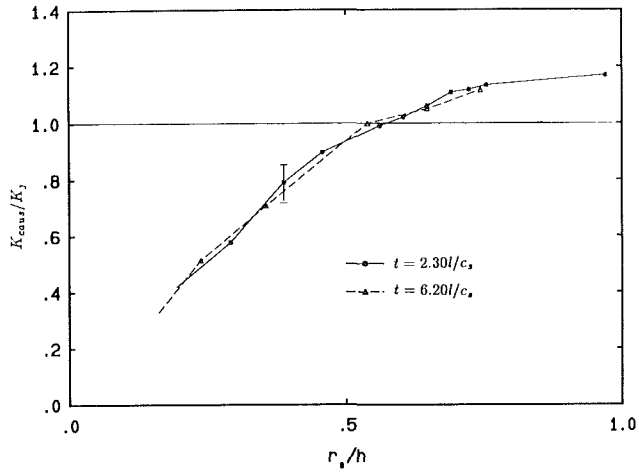
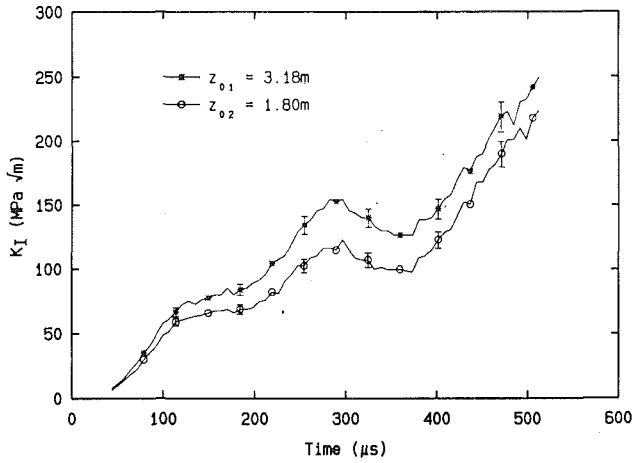


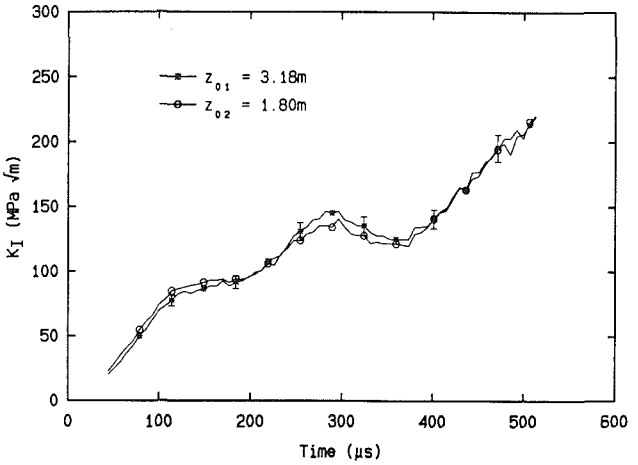
Fig. 10(*b*) Plot of K_{caus}/K_j versus r_0/h from the three-dimensional elastodynamic simulation

As a parenthetical note, it is interesting that the variation in K_{I1}^d and K_{I2}^d—the two values for the dynamic stress intensity factor as measured from pairs of bifocal caustics—turns out to be in quite good *quantitative* agreement with the results of Fig. 10(*b*). That is, the ratio of K_{I1}^d to K_{I2}^d as obtained from experiments seem to be in close agreement with that obtained from Fig. 10(*b*) for the corresponding initial curve radii. This can be seen in Fig. 11 where the numerically generated results of Fig. 10(*b*) are used to "scale" the results of one particular experiment. To do this, *assume* that the results of Fig. 10(*b*) (K_{caus}/K_j versus r_0/h) hold for the whole duration of the loading. The experimental data (corresponding to the two z_0s) shown in Fig. 11(*a*) can then be scaled, for each time, to the corre-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



(a)



(b)

Fig. 11 (*a*) Raw and (*b*) "scaled" experimental data for specimen (3*q*)

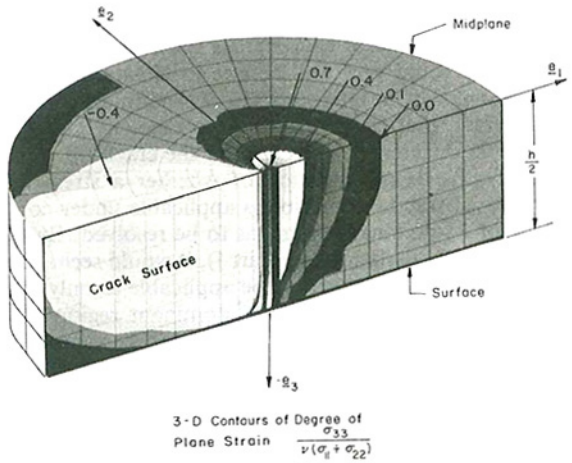


Fig. 12 Representative near-tip variation of the degree of plane strain

sponding "K_J" value by means of Fig. 10(*b*). The resulting "scaled" dynamic stress intensity factor histories corresponding to the two z_0s are shown in Fig. 11(*b*). It is seen that by this procedure the deviation between the K_I^d-values observed in the experiments from pairs of bifocal caustics essentially reduce to within expected experimental scatter for the whole duration of loading. Thus, it would appear that the main

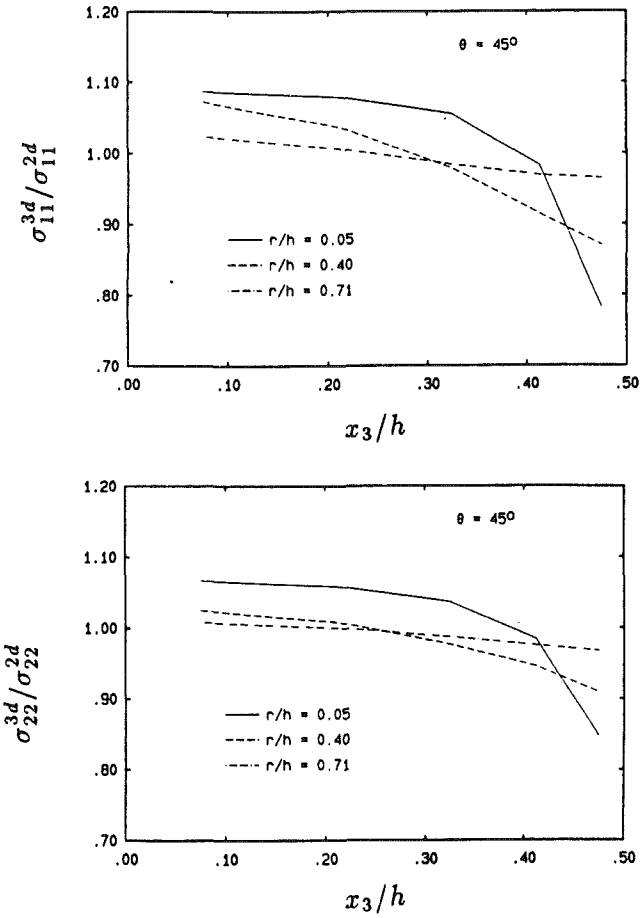


Fig. 13 Representative through thickness variation of normalized stress components

reasons for the observed experimental variation are accounted for by the three-dimensional elastodynamic simulation. Note that it is purely to highlight this point that the above "scaling" procedure was adopted. It is not the intention here to offer Fig. 10(*b*) as some kind of an empirical "correction curve" for experimental data.

Finally, it is instructive to look into some additional features of the near-tip field quantities. Figure 12 shows the near-tip variation of the degree of plane strain for one typical time in the simulation. In regions where plane stress is a good approximation to the actual three-dimensional state of stress, the degree of plane strain must be close to zero. It is seen that, by this measure, plane-stress conditions are obtained at radial distances greater than about one-half plate thickness from the crack front. Also note that the radial extent of the region of three-dimensionality is not uniform around the crack front but varies substantially with angle.

The thickness variations of stress components (σ_{11} and σ_{22} normalized by the corresponding full-field two-dimensional values) are shown for one representative time in Figs. 13(*a*), 13(*b*) for a range of radial distances from the crack tip along the $\theta = 45$ deg line. It can be seen that the deviations in stresses from the two-dimensional fields are largest toward the crack tip and the free surface. Also, these plots underscore the point that the assumption of plane stress is indeed a good one for regions of radial extent greater than about one-half plate thickness away from the crack front. Note, however, that this does not imply that the asymptotic K_I^d-field is dominant in the plane-stress region outside the near-tip zone of three-dimensionality.

4 Discussion

Based on two and three-dimensional elastodynamic simu-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

lations of the drop-weight tests, the following conclusions can be drawn:

(i) The dynamic asymptotic field, while sufficiently accurate for $r/h \to 0$, becomes increasingly inadequate for larger radial distances even in a purely two-dimensional setting.

(ii) The three-dimensional nature of the dynamic crack-tip field, which is seen to be confined to within at most one-plate thickness radial extent around the crack-tip, exhibits largest deviation from the full-field plane-stress results for $r/h \to 0$.

(iii) The above two results together imply that the three-dimensional structure of the near-tip surface, coupled with the transient nature of the local fields, appear to preclude any sizeable region of K_I^d-dominance around the crack tip.

The experiments reported in Part I had also indicated the lack of an underlying K_I^d-dominant field. This was observed experimentally for both the dynamically loaded stationary crack as well as for rapidly propagating cracks. While it is recognized that the results of this work are specific to the configuration studied, it is not inconceivable, especially in view of lack of countervailing evidence, that a similar result could hold in other settings. It might thus be instructive to speculate on the validity of the presumed K_I^d-dominance in some of the experiments reported in the literature.

Some Implications of Lack of K_I^d-Dominance. As far as the dominance of a K_I-field is concerned, the ratio of the smallest pertinent in-plane length to the specimen thickness can be thought of as a relevant geometry parameter. As pointed out in Part I, the possible extent of a K_I-dominant annulus around the crack-tip is bounded from within by the maximum extent of (a) the process zone, (b) the nonlinear region, and (c) the three-dimensional region. For nominally brittle materials such as those considered in this work, it appears that three-dimensionality is the most critical of the three. Also, the outer bound for a K_I-dominant annulus around the crack-tip is expected to depend on a relevant in-plane length scale. Thus, under static conditions at least, it is expected that the ratio a/h (where "a" is the smallest relevant in-plane length and "h" is the specimen thickness) must be sufficiently large for a K_I-field to survive the three-dimensional region and establish its dominance over some finite domain. In dynamic problems, the issue is much more involved with additional requirements dictated by the nature of the loading and the time required for stress-wave information to reach regions outside the three-dimensional zone (see, for example, Ravi-Chandar and Knauss, 1987). These, however, would seem to only further restrict the possibility of obtaining a K_I^d-dominant region in a real experiment. Thus, if a/h fails to be sufficiently large, it is not expected that a K_I^d-dominant field would prevail over any finite domain around the crack tip. In particular, note that in the experiments reported in this work, a/h is about 9 (based on the uncracked ligament as the relevant in-plane length) and, for this value, a K_I^d-dominant field was not observed. It might thus be expected that the existence of a K_I^d-dominant region is not assured for specimens for which a/h is of the same order as this work. It is not the authors' intention here to make a detailed review of the experimental literature with a view to ascertaining whether K_I^d-dominance prevailed in these experiments or not. Rather, the parameter a/h is suggested here as one measure by which the interested reader might gauge the relevance of the results of this work to other experiments. In the following, some issues on which the results of this work might have a bearing are briefly discussed.

On K_{ID}—$\dot{a}$ *Relations:* The question of the precise functional dependence of the dynamic fracture toughness on the crack velocity has appropriated a substantial share of the resources devoted to the study of dynamic fracture. The argument, for the most part, has centered around the issue of uniqueness of relationship between K_{ID} and $\dot{a}$. Kobayashi and Dally (1980), Rosakis, Duffy, and Freund (1984), and Zehnder and Rosakis (1989) provide data sets that seem to indicate that the K_{ID}—$\dot{a}$ relation is indeed (to within experimental error) a unique material property. The results of Kobayashi and Mall (1978) and Ravi-Chandar and Knauss (1984)—using photoelasticity and the method of caustics in transmission, respectively—suggest, however, that there is no such one-to-one correspondence. Is the substantial scatter in these data sets due to a lack of a unique correspondence between K_{ID} and $\dot{a}$, or due to a more fundamental lack of a K_I^d-dominant field in the experiments? Also, in all these experiments, the question of whether it is indeed the dynamic stress intensity factor that is being measured needs to be looked into.

On Specimen Dependence of K_{ID}: By far the most troubling of all the experimental results are those that indicate that the dynamic fracture toughness could be specimen dependent (Kalthoff, 1983; Dahlberg, Nilsson, and Brickstad, 1985). These results, rather than merely contending about the parametric dependence of a material property—the fracture toughness—would, if true, seem to indicate that the stress intensity factor based fracture criterion is fundamentally flawed. However, it is possible to attribute the apparent observed specimen dependence of K_{ID} to specimen-dependent differences in the near-tip field (i.e., the near-tip fields, not being K_I^d-dominant, might have a different structure from specimen to specimen). Moreover, as has been shown in this work, much larger differences (than the 20 percent or so that is seen in the results of Kalthoff (1983)) could possibly have been obtained for the *same* type of specimen if, in the caustics experiments used to obtain these results, a different range of initial curves had been used by changing the specimen to image plane distance z_0.

On K_{ID}—$\ddot{a}$ *Relations:* Attempts by Takahashi and Arakawa (1987)—using the method of caustics again—to show acceleration dependence of the dynamic fracture toughness can also be deemed inconclusive for precisely the same reasons as above.

On Photo-elasticity versus Caustics: Nigam and Shukla (1988) have recently tried to compare the methods of photoelasticity and transmission caustics by doing tests on identical specimens under identical loading. They show that while both methods work well for static problems, the method of photoelasticity gives values for the dynamic stress intensity factor which are about 30–50 percent higher than those obtained through the method of caustics. As shown in this paper, this discrepancy could very well be due to violation of K_I^d-dominance assumed in the interpretation of experimental data. Note, in particular, that caustics and photo-elasticity data typically come from different regions around the crack tip.

Finally, the larger question of whether a stress intensity factor based fracture criterion is applicable under conditions of lack of K_I^d-dominance remains to be resolved. By the very nature of its formulation (see Part I), it would seem that a K_I^d-based fracture criterion would be applicable to only that class of problems where there is a K_I^d-dominant region confining the near-tip nonlinear and fracture process zones. Thus, it could be argued that investigations designed to resolve questions of the dependence of the dynamic fracture toughness on crack velocity, acceleration, or specimen geometry, etc., would be meaningful only if (i) the experimental configuration utilized is shown to lead to a region of K_I^d-dominance and (ii) the observed variations are significant in comparison to the uncertainty associated with the assumptions of the experimental techniques used.

5 Conclusion

In this paper, as well as in Part I, an attempt has been made to investigate the consequences of interpreting experimental results using the method of reflected caustics under the assumption of K_I^d-dominance. It was shown that this assumption

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

could lead to substantial errors in the measured results. It is our view that many of the apparent discrepancies in the experimental literature might arise from the lack of an underlying K_I^d-dominant or higher-order two-dimensional region assumed in the interpretation of the experimental data. The assumption that the near-tip region is K_I^d-dominant (or even two-dimensional, for that matter) is only an approximation and, as such, entails a corresponding uncertainty in the measured values for the dynamic stress intensity factor. We feel that it is essential to undertake studies in the spirit of the one reported here to investigate the assumptions of other experimental techniques (photo-elasticity, transmission caustics, etc.) and evaluate the costs incurred in terms of loss of accuracy.

Acknowledgments

The authors would like to thank Prof. M. Ortiz for providing a version of the finite element program, FEAP. Support of the Office of Naval Research through contract N00014-85-K-0599 is gratefully acknowledged. The computations were performed using the facilities of the San Diego Supercomputer Center and were made possible through an NSF-PYI grant MSM-84-51204 to the second author.

References

Bathe, K., 1982, *Finite Element Procedures in Engineering Analysis*, Prentice Hall.

Dahlberg, L., Nilsson, F., and Brickstad, B., 1980, "Influence of Specimen Geometry on Crack Propagation and Arrest Toughness," *Crack Arrest Methodology and Applications*, G. T. Hahn et al., ed., ASTM-STP 711, pp. 89–108.

Dally, J. W., Fourney, W. L., and Irwin, G. R., 1985, "On the Uniqueness of K_{ID}—$\dot{a}$ Relation," *International Journal of Fracture*, Vol. 27, pp. 159–168.

Hinton, E., and Campbell, J. S., 1974, "Local and Global Smoothing of Discontinuous Finite Element Functions Using a Least Squares Method," *International Journal for Numerical Methods in Engineering*, Vol. 8, pp. 461–480.

Hughes, T. J. R., and Belytschko, T., 1983, "A Precis of Developments in Computational Methods for Transient Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1033–1041.

Kalthoff, J. F., 1983, "On Some Current Problems in Experimental Fracture Mechanics," *Workshop on Dynamic Fracture*, W. G. Knauss et al., ed., California Institute of Technology, pp. 11–35.

Kobayashi, A. S., and Mall, S., 1978, "Dynamic Fracture Toughness of Homalite 100," *Experimental Mechanics*, Vol. 18, pp. 11–18.

Kobayashi, T., and Dally, J. W., 1980, "Dynamic Photo-elastic Determination of the $\dot{a}$—K Relation for the 4340 Steel," *Crack Arrest Methodology*, G. T. Hahn et al., ed., ASTM-STP 711, pp. 89–108.

Krishnaswamy, S., 1988, "On the Domain of Dominance of the Asymptotic Elastodynamic Crack-Tip Fields," doctoral dissertation, California Institute of Technology, Pasadena.

Levy, N., Marcal, P. V., and Rice, J. R., 1971, "Progress in Three-Dimensional Elastic Stress Analysis for Fracture Mechanics," *Nuclear Engineering and Design*, Vol. 17, pp. 64–75.

Nakamura, T., and Parks, D. M., 1987, "Three-Dimensional Field Near the Crack Front of a Thin Elastic Plate," presented at the 24th Annual Meeting of the Society of Engineering Science, Sept. 21–23, 1987, Salt Lake City, Utah.

Nakamura, T., and Parks, D. M., 1988, "Conditions for J-Dominance in Three-Dimensional Thin Cracked Plates," *Analytical, Numerical and Experimental Aspects of Three-Dimensional Fracture Process*, A. J. Rosakis et al., ed., ASME-AMD, Vol. 91, pp. 227–238.

Nakamura, T., Shih, C. F., and Freund, L. B., 1986, "Three-Dimensional Transient Analysis of a Dynamically Loaded Three-Point Bend Ductile Fracture Specimen," Brown University Report, ONR 0365/13, Brown University.

Narasimhan, R., and Rosakis, A. J, 1988, "Three-Dimensional Effects Near a Crack-Tip in a Ductile Three-Point Bend Specimen; Part I: A Numerical Investigation," SM-Report 88-6, GALCIT, California Institute of Technology.

Nigam, H., and Shukla, A., 1988, "Comparison of the Techniques of Transmitted Caustics and Photoelasticity as Applied to Fracture," *Experimental Mechanics*, Vol. 28, No. 2, pp. 123–133.

Parsons, I. D., Hall, J. F., and Rosakis, A. J., 1986, "A Finite Element Investigation of the Elastostatic State Near a Three-Dimensional Edge Crack," GALCIT Report SM 86-29, California Institute of Technology.

Ravi-Chandar, K., and Knauss, W. G., 1984, *International Journal of Fracture*, Vol. 25, pp. 247–262.

Ravi-Chandar, K., and Knauss, W. G., 1987, "On the Characterization of the Transient Stress Field Near the Tip of a Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 72–78.

Rosakis, A. J., and Freund, L. B., 1981, "The Effect of Crack-Tip Plasticity on the Determination of the Dynamic Stress-Intensity Factor by the Optical Method of Caustics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 302–308.

Rosakis, A. J., Duffy, J., and Freund, L. B., 1984, "The Determination of Dynamic Fracture Toughness of AISI 4340 Steel by the Shadow Spot Method," *Journal of the Mechanics and Physics of Solids*, Vol. 32, pp. 443–460.

Smith, R. H., and Freund, L. B., 1988, "Three-Dimensional Finite Element Analysis of Steady Elastodynamic Crack Growth in a Plate," presented at the joint ASME-SES Conference, June 1988, Berkeley, Calif.

Takahashi, K., and Arakawa, K., 1987, "Dependence of Crack Acceleration on the Dynamic Stress-Intensity Factor in Polymers," *Experimental Mechanics*, Vol. 27, No. 2, pp. 195–199.

Taylor, R. L., 1977, "Computer Procedures for Finite Element Analysis," *The Finite Element Method*, 3rd ed., by O. C. Zienkiwiecz, McGraw-Hill, New York.

Zehnder, A. T., and Rosakis, A. J., 1989, "Dynamic Fracture Initiation and Propagation in 4340 Steel Under Impact Loading," *International Journal of Fracture*, to appear.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

C. Betegón

Department of Construction,
University of Oviedo,
Oviedo, Spain

J. W. Hancock

Department of Mechanical Engineering,
University of Glasgow,
Glasgow, Scotland

Two-Parameter Characterization of Elastic-Plastic Crack-Tip Fields

Plane-strain elastic-plastic crack-tip fields have been modeled with modified boundary layer formulations based on the first two terms K *and* T, *of the elastic field. These formulations match the appropriate full field solutions. Compressive* T *stresses reduce the stresses by an amount which is independent of radial distance, corresponding to the introduction of a second term in addition to the dominant plastic singularity. Geometries which maintain* J*-dominance are characterized by zero or positive* T *stresses, while geometries with negative* T *stresses can be described by a two-parameter characterization using* J *and* T *into full plasticity.*

Introduction

The characterization of crack-tip stress and strain fields is fundamental to fracture mechanics. In classical linear elastic fracture mechanics, the appropriate parameter is the stress intensity factor K, whose significance can be seen by expanding the stress field in cylindrical coordinates (r,θ) about the crack tip, following the work of Williams (1957).

$$\sigma_{ij} = A_{ij}(\theta) r^{-1/2} + B_{ij}(\theta) + C_{ij}(\theta) r^{-1/2} + \dots \tag{1}$$

The first term in the expansion is singular at the crack tip, whereas the remaining terms are finite and bounded. Elastic fracture mechanics is thus based on the premise that fracture processes which occur close to the crack tip are only determined by the first term in the expansion, allowing the asymptotic elastic stress field of a symmetrically loaded mode I crack to be expressed in the familiar form:

$$\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta). \tag{2}$$

Larsson and Carlsson (1975) have, however, demonstrated that the second term in the series has a significant effect on the shape and size of the plastic zone which develops at the crack tip. In the notation of Rice (1974), the second term in the expansion is denoted the T stress and can be regarded as a stress parallel to the crack flanks.

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \frac{K}{\sqrt{2\pi r}} \begin{bmatrix} f_{11}(\theta) & f_{12}(\theta) \\ f_{21}(\theta) & f_{22}(\theta) \end{bmatrix} + \begin{bmatrix} T & 0 \\ 0 & 0 \end{bmatrix} \tag{3}$$

The magnitude of the T stress is defined through a biaxiality parameter B, introduced by Leevers and Radon (1983)

$$B = \frac{T\sqrt{\pi a}}{K}. \tag{4}$$

Although the stress intensity factor K has been tabulated for a wide range of geometries, the biaxiality parameter B is only available for a more limited number of cases (Leevers and Radon, 1983; Cardew et al., 1983; Kfouri, 1986; James-Watt, 1989). The evolution of elastic-plastic fracture mechanics has also been based on the development of suitable characterising parameters, notably the J-integral, Rice (1968). Hutchinson (1968 a,b) and Rice and Rosengren (1968), (henceforth HRR), have analyzed the asymptotic crack-tip fields using a variational technique in which the field was described only by a single dominant term close to the crack tip for materials showing a nonlinear stress-strain response such as may be represented by a Ramberg-Osgood power law.

$$\frac{\epsilon}{\epsilon_0} = \frac{\sigma}{\sigma_0} + \alpha \left(\frac{\sigma}{\sigma_0}\right)^n \sigma \geq \sigma_0. \tag{5}$$

On this basis the stress and strain fields under mode I deformation can be regarded as a series expansion about the crack tip, in which the leading term has the form

$$\sigma_{ij} = \sigma_0 \left[\frac{J}{\epsilon_0 \sigma_0 \alpha I_n r}\right]^{1/n+1} \tilde{\sigma}_{ij}(\theta, n) \tag{6}$$

$$\epsilon_{ij} = \frac{\sigma_0 \alpha}{E} \left[\frac{J}{\epsilon_0 \sigma_0 \alpha I_n r}\right]^{n/n+1} \tilde{\epsilon}_{ij}(\theta, n) \tag{7}$$

where $\tilde{\sigma}_{ij}(\theta,n)$, $\tilde{\epsilon}_{ij}(\theta,n)$ and I_n are tabulated functions of the strain-hardening exponent n, and where appropriate the angle θ. The HRR field is essentially a small geometry change solution, in which the crack tip is assumed to remain sharp. Equivalently, the solution can be regarded as being appropriate at distances from the crack tip at which the effects of crack-tip blunting can be neglected. This is usually taken to be two or three crack-tip openings, or equivalently distances of the order of $2J/\sigma_0$. Within this distance it is necessary to adopt a large strain formulation to investigate the near-tip blunting region.

McMeeking (1977) has analyzed this problem using a boundary layer formulation in which the remote elastic field was modeled by tractions corresponding only to the K term of the elastic field. In the current notation this corresponds to a

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 27, 1989; final revision, Dec. 20, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

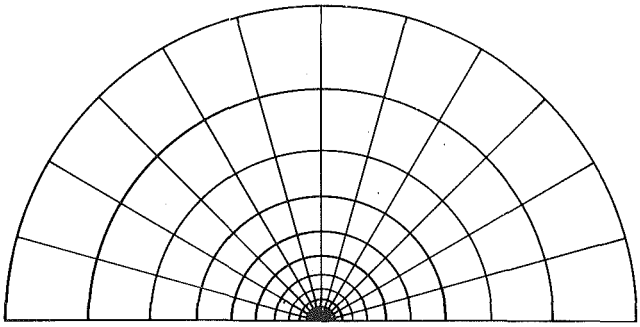


Fig. 1 The mesh of the boundary layer formulation

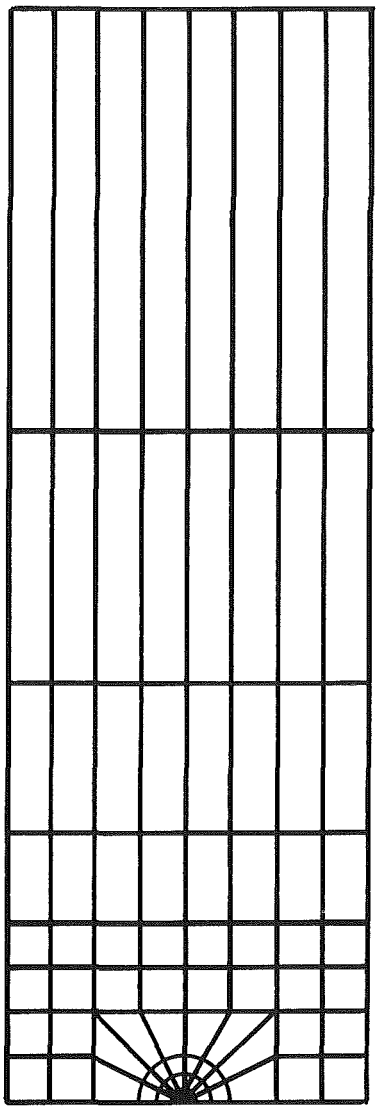


Fig. 2 The mesh of a symmetric quarter of a center-cracked panel (a/W) = 0.5 subject to displacement-controlled tensile boundary conditions

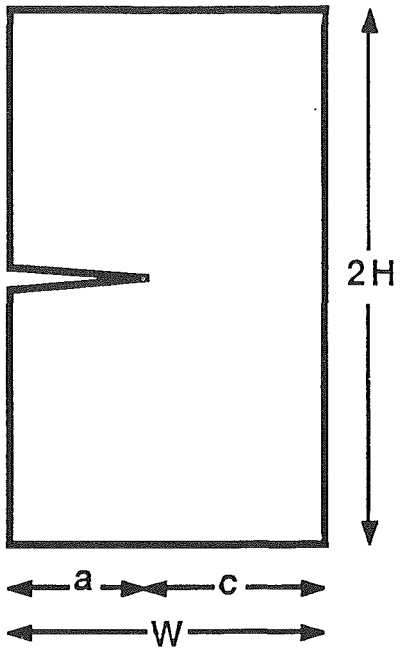


Fig. 3 The geometry of the edge-cracked bars. For the SENB geometry $a/W = 0.3$ and 0.9, $H/w = 3$

biaxiality parameter, $B = 0$ in equation (4). More recently Bilby, Cardew, Goldthorpe and Howard (1986) have examined the effect of the second term in the elastic expansion using a modified boundary layer formulation incorporating both K and T terms in the boundary conditions. These calculations show the sensitivity of the blunting field to negative, that is compressive T stresses, which lower the maximum stress ahead of the blunting crack.

In nonhardening plasticity, it is known that the fully plastic deformation field is not unique, and that the appropriate slip line field depends both on the geometry of the body and the way that it is loaded (McClintock 1968). For center-cracked panels, subject to tension, this leads to a pronounced change in the nature of crack-tip deformation between small-scale yielding and full plasticity. The conditions under which the crack-tip deformation maintains the chararacter of the small-scale yielding field has been defined as J dominance, and quantitatively studied by McMeeking and Parks (1979) and Shih and German (1981). In the first work, a contained yielding solution was modeled by a boundary formulation with no T stress, and compared with full-field solutions of deeply edge cracked bars in tension and bending. McMeeking and Parks (1979) concluded that the deformation fields were similar as long as the uncracked ligament was greater than $200J/\sigma_0$ in tension, and $25J/\sigma_0$ is bending. Similar calculations by Al-Ani and Hancock (1991) have shown that for short cracks in bending and tension J dominance has been lost before $J = a\sigma_0/200$.

An alternative approach, reinforcing the findings of McMeeking and Parks (1979) has been presented by Shih and German (1981) who compared the full-field small geometry change solutions with the HRR field, with the criterion that the calculated stress in the full-field solution should be within ten percent of the HRR field at a distance distance $2J/\sigma_0$. In the present work, the approach developed by Shih and German (1981) has been adopted to investigate the effect of elastic T stresses on J dominance. The crack-tip field has been firstly modeled by plane-strain elastic-plastic boundary layer formulations involving both the K and T terms. These have then been correlated with full-field solutions for geometries which exhibit positive and negative T stresses. This has finally led to a consideration of the conditions for J dominance within elastic-plastic flow fields.

Numerical Methods

Plane-strain crack-tip deformation was modeled by boundary layer formulations using focused meshes of the type shown in Fig. 1. The meshes typically involved 240 eight-noded hybrid isoparametric elements (Hibbitt, Karlsson, and Sorensen, 1984b) consisting of 20 rings and 12 elements concentric with the crack tip. The crack tip thus comprised 25 independent, but initially coincident nodes. Mesh refinement was such that

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

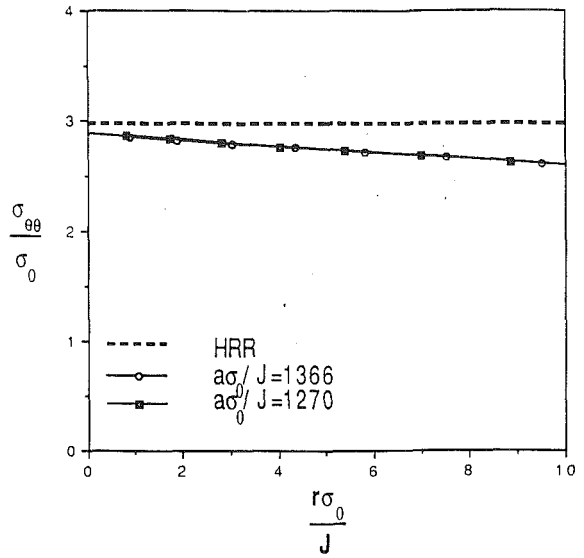


Fig. 4 The tangential stress directly ahead of the crack in an elastic nonhardening plastic boundary layer formulation, $B=0$, $n=\infty$

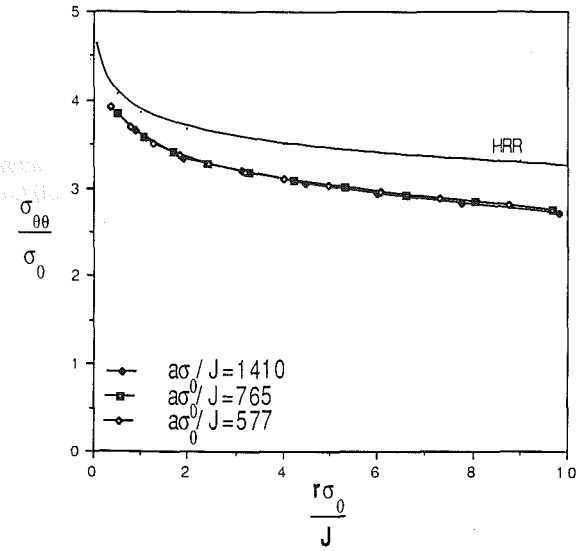


Fig. 5 The tangential stress directly ahead of the crack in the boundary layer formulation, $B=0$, $n=13$

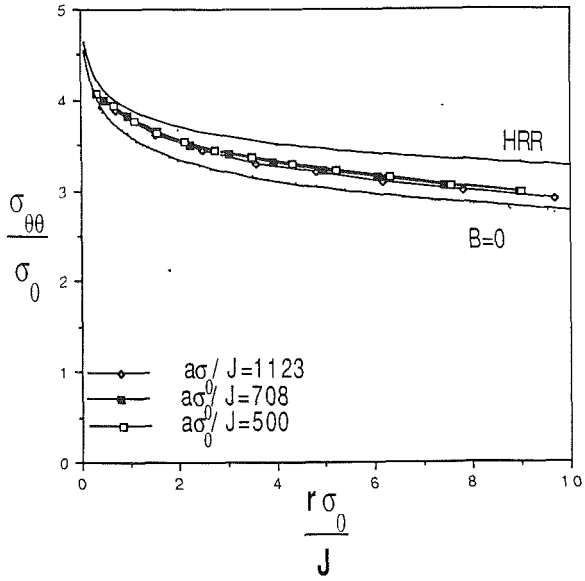


Fig. 6 The tangential stress directly ahead of the crack in a modified boundary layer formulation, $B=+1.06$, $n=13$

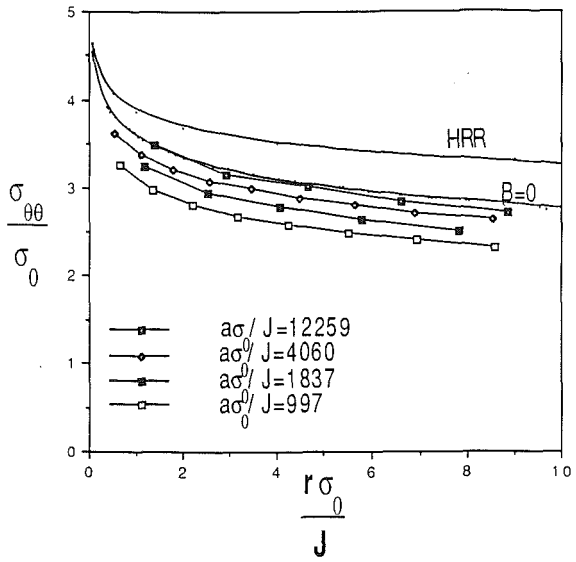


Fig. 7 The tangential stress directly ahead of the crack in a modified boundary layer formulation, $B=-1.06$, $n=13$

the radius of the first ring of elements was less than one-thousandth of the radius of the outer ring. Displacement boundary conditions were imposed on the outer boundary corresponding to the displacements associated with a K field plus the displacements due to the T stress.

Although the formulation does not involve an explicit crack or ligament length, a dimensional scale is introduced by the radius at which displacements appropriate to a given K and T field are applied. The K and T terms were increased in a proportional way by an imposed biaxiality parameter B. Solutions were obtained for plastic zone radii very much less than the radius of the outer boundary in accord with the procedure described by Tracey (1976) and Parks and Wang (1988). For plastic zones within these limits, the $B=0$ solutions satisfied the necessary condition for a small-scale yielding problem, being self-similar when the stresses were nondimensionalized by the yield stress σ_0, and the distances ahead of the crack tip were nondimensionalized by J/σ_0.

In the boundary layer formulations, J was determined both by the virtual crack extension method of Parks (1977), as modified by Li, Shih, and Needleman (1986), and implemented in ABAQUS; and from the K component of the outer field. For purely elastic solutions, the two methods gave the same results. Similarly, for elastic-plastic solutions with $T=0$, the two methods gave identical results. However, the presence of a T stress in the elastic-plastic calculations introduced a small difference, also noted by Bilby et al. (1987), such that the virtual crack extension method gave slightly lower results than the K field for negative T stresses but larger results for positive T stresses. As an example at $T/\sigma_0=0.25$ in the nonhardening solutions the difference was 0.4 percent while at $T/\sigma_0=-0.5$, the difference was 5 percent. This effect may either arise from a lack of path independence, which is not guaranteed for flow plasticity or from the finite size of the mesh. However, this was not considered to be a major effect, and in the data presented, J was always determined by virtual crack extension. The T stress, as an essentially elastic concept, was calculated from the remote K field using the biaxiality parameter B. In the full-field solutions with extensive plasticity the same procedure was adopted, and this essentially involves identifying T with the elastic component of J.

The material response was described by a Ramberg-Osgood power law of the form given in equation (5), with the exponent

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

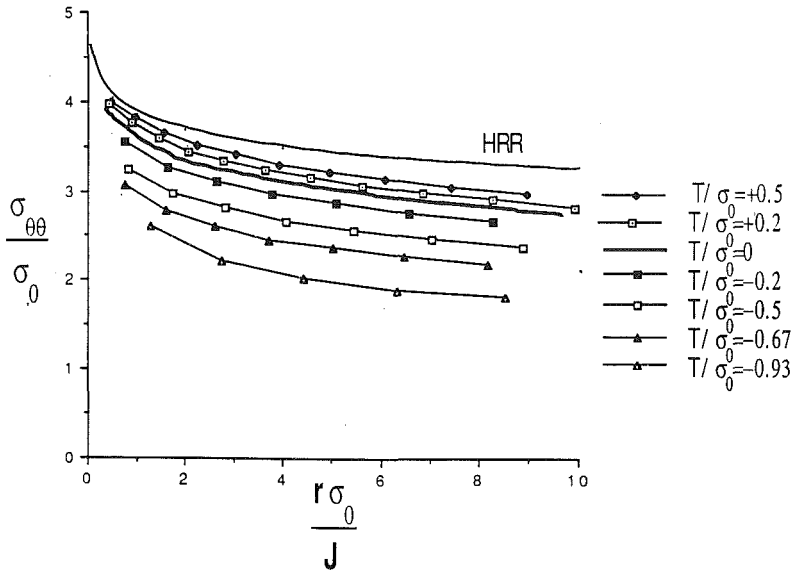


Fig. 8 The tangential stress field directly ahead of a cracks in boundary layer formulations at different levels of the T stress, $n = 13$

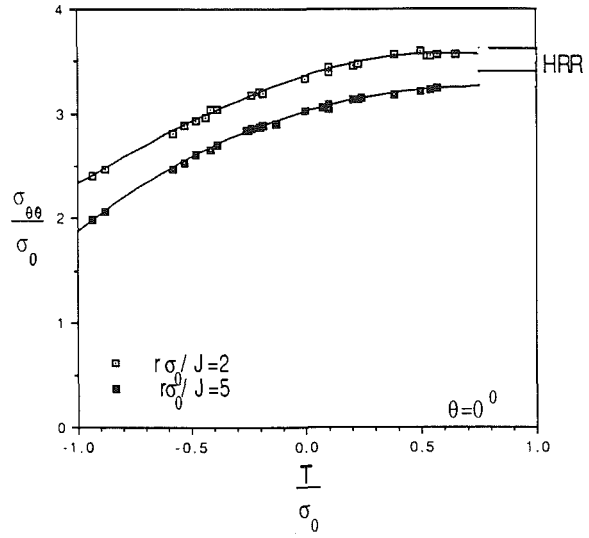


Fig. 9 The tangential stress directly ahead of a cracks at distances $2J/\sigma_0$ and $5J/\sigma_0$ as a function of T/σ_0, $n = 13$

$n = 3, 13$, and infinity. Poisson's ratio was set as 0.3, and $\alpha = 3/7$, while the ratio of the yield stress σ_0 to the elastic modulus was 0.002. The analyses were based on small-strain theory flow plasticity (Hibbitt, Karlsson, and Sorensen, 1984a), adopting an incremental form of the Prandtl-Reuss flow rule. Tangential stresses were obtained both directly ahead of the crack ($\theta = 0$) and at ($\theta = \pi/4$) by extrapolating the stress field to the nodes.

The three geometries examined with full-field solutions are shown in Figs. 2 and 3. These geometries were chosen to represent a wide range of T stresses. The center-cracked tensile panel, $a/W = 0.5$, has a biaxiality parameter $B = -1.06$, (Leevers and Radon, 1983). In contrast, a deeply edge-cracked bend bar, $a/W = 0.9$, was chosen to be representative of geometries with positive T stresses. For this geometry, B is $+1.1$ (Al-Ani and Hancock, 1991). The final specimen was an edge-cracked bend bar (SENB), with an(a/W) ratio of 0.3, under pure bending, which has a B value close to zero (Al-Ani and Hancock, 1991). In all cases the crack was modeled with a focused mesh with independent coincident nodes using eight-noded isoparametric plane-strain elements. The center-cracked panel (henceforth CCP) was modeled by 208 plane-strain second-order elements giving a problem size of approximately 1400 degrees-of-freedom, which was solved on a VAX 11/750 computer. The deeply edge-cracked bend bar was modeled with 191 elements giving a problem with 1288 degrees-of-freedom. The CCP specimen was subject to uniform tensile displacement boundary conditions. Both bend bars were subject to force loading conditions corresponding to a pure couple on the remote boundary.

Results

Boundary Layer Formulations. Figures 4 and 5 show the tangential stress $\sigma_{\theta\theta}$ directly ahead of the crack ($\theta = 0$) for hardening and nonhardening boundary layer formulations in which the T stress is zero ($B = 0$). The stresses are normalized by the initial yield stress σ_0, while the radial distances of a point from the crack tip, r, are nondimensionalized by J/σ_0. The data is self-similar in the sense that data obtained for a given applied K field falls on the same curve as that for a higher K, but falls below the HRR field determined from the tabulated constants of Hutchinson (1968) and Shih (1983).

Figure 6 illustrates the behavior of the stresses on the same plane with a positive biaxiality parameter $B = +1.06$ which produces tensile T stresses which increase with deformation. The data are again compared both with the HRR and $B = 0$ fields. In this case the data are not quite self-similar. Early in the deformation, the stresses are comparable to the $B = 0$ field, but with increased loading the stresses rise slightly, until the data closely approximate to the HRR field.

Figure 7 illustrates the effect of a negative biaxiality parameter B which produces a compressive T stress. In this case the stresses are again initially close to the $B = 0$ data for small T stresses, but as the T stress becomes more negative, the stresses fall and depart significantly from the HRR field.

The data exemplified in Figs. 5, 6, and 7 are combined with other data and presented in Fig. 8 as a family of curves which are functions of the T stress, but independent of the biaxiality, B. Particular attention is paid to the development of the stress ahead of the crack at distances $r = 2J/\sigma_0$ and $5J/\sigma_0$, as illustrated in Fig. 9, while the effect of hardening is shown in Fig. 10. The data approach the HRR field for positive T stresses, but fall markedly below the HRR field with negative T stresses.

Full-Field Solutions. Three complete cracked geometries have been studied with full-field solutions. The geometries were

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

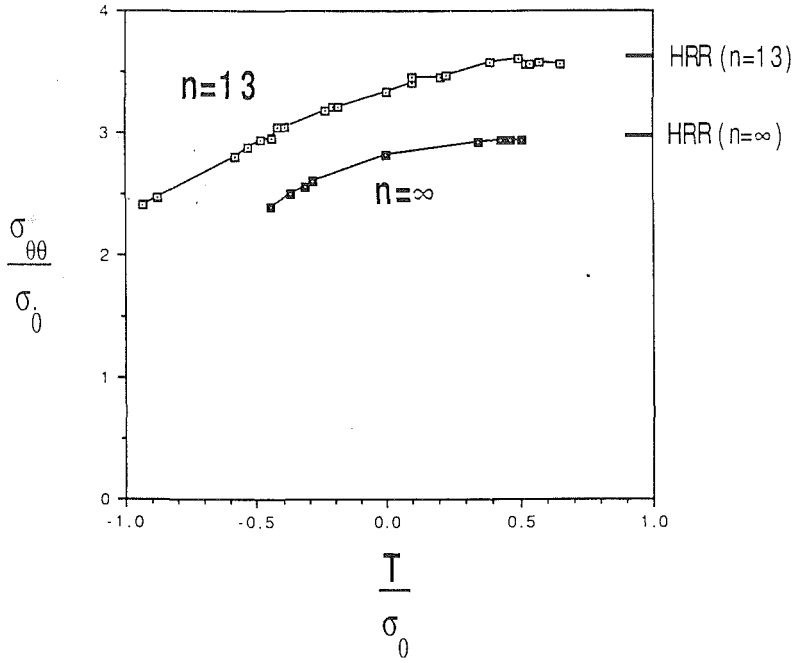


Fig. 10 The tangential stress directly ahead of the crack at distances $2J/\sigma_0$ for two strain-hardening rates, $n=13$ and $n=\infty$

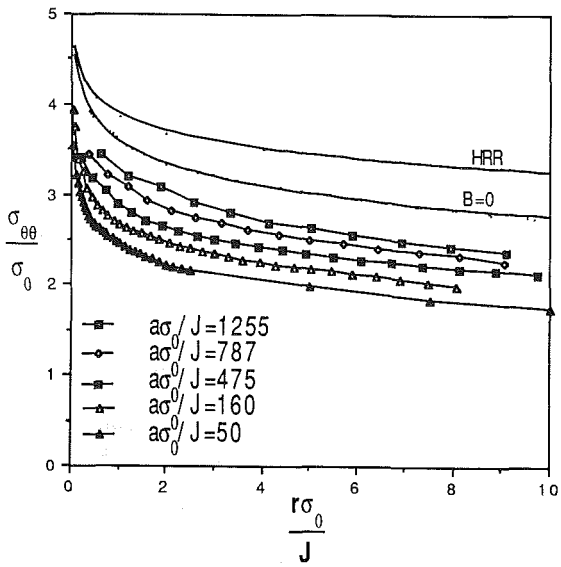


Fig. 11 The tangential stress directly ahead of the crack in central crack panel under uni-axial tension, $n=13$

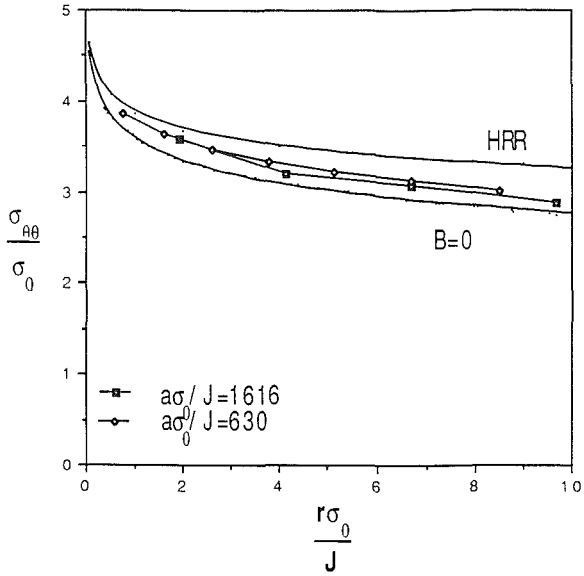


Fig. 12 The tangential stress directly ahead of the crack in a single-edge notched bend bar (SENB, $a/W=0.9$, $n=13$)

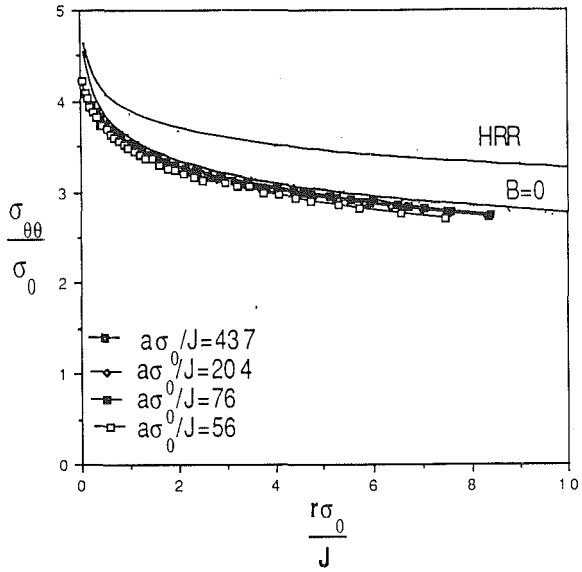
chosen to be representative of positive, zero, and negative T stresses. The tangential stess straight ahead of the CCP crack, which has negative T stresses, are shown in Fig. 11. Even at the smallest loading for which the plastic zone is large enough to be studied, the stresses fall significantly below the HRR field and continue to fall with increasing deformation both at $\theta=0$ and $\theta=\pi/4$. The deeply edge-cracked bend bar subject to a pure couple on the remote boundary is shown in Fig. 4. The biaxiality ratio for this geometry is approximately $+1.1$. The stresses ahead of the crack again start close to the $B=0$ solution but rise slightly with increasing deformation until they approach the HRR field as shown in Fig. 12. Finally, the corresponding data for the edge-cracked bend bar with a biaxiality parameter close to zero is shown in Fig. 13.

The boundary layer formulations are compared with a full-field solution of a CCP geometry ($B=-1.06$), by examining the stresses at distances $2J/\sigma_0$ and $5J/\sigma_0$ from the crack tip, as shown in Fig. 14 for $n=13$. The two-term boundary layer formulation is shown to match the full-field solution even at deformation levels of $a\sigma_0/J=50$ which corresponds to full plasticity. Similarly, for the $B=0$ problem (SENB, $a/W=0.3$), the full-field solution shown in Fig. 13 closely matches the corresponding boundary layer formulation in which only the K term is applied on the boundaries. Finally, the same match between the two-term boundary layer formulation and the full-field solution was observed for the deeply cracked bend bar (SENB, $a/W=0.9$) which exhibited stresses close to the HRR field in confirmation of the calculations of Shih and German (1981).

Discussion

One-term boundary layer formulations ($B=0$) give rise to self-similar stress fields, which although being characterized by J, differ from the HRR field. As an example, the $B=0$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



[Fig. 13 The tangential stress directly ahead of the crack in a single-edge notched bend bar (SENB, $a/W = 0.3$, n = 13), after Al-Ani and Hancock (1989b)

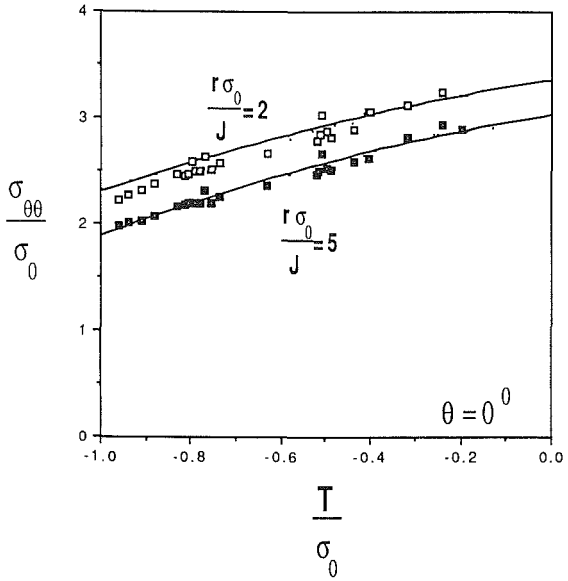


Fig. 14 A correlation between the full-field solution for a CCP geometry ($B = -1.06$) and the boundary layer solution at distances $2J/\sigma_0$ and $5J/\sigma_0$, $n = 13$. The full-field solution is given by the data points and the boundary layer formulation by the solid line.

field for the nonhardening solution at the crack tip is three percent lower than the Prandtl value of $2.97\sigma_0$ which is the nonhardening limit of the HRR field (Hutchinson, 1968b). For $n = 13$ the numerical solution is eight percent lower than the HRR field at a distance $2J/\sigma_0$, and the discrepancy increases with hardening rate. These results are consistent with the original calculations of Tracey (1976), and those more recently presented by Parks and Wang (1988). It is necessary to conclude that for low hardening rates the HRR field dominates at distances less than $2J/\sigma_0$, which is within the region which is most appropriately analysed by large geometry change solutions. In these conditions the field, surrounding the blunting zone, close to the crack tip, is better described by the $B = 0$ field in the small-scale yielding limit in which the applied stress intensity approaches zero.

When a T stress is introduced, tensile T stresses cause the stress field to approach the HRR field, while compressive T stresses cause the stresses to fall significantly. The effect of

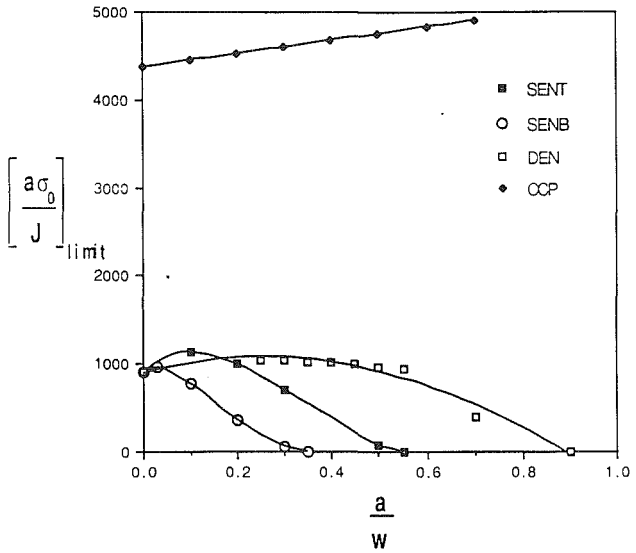


Fig. 15 The size requirements for J dominance for a range of geometries as a function of the crack depth (a/w)

negative T stresses is to reduce the stress level ahead of the crack by an amount that depends only on T, and is independent of the distance $r\sigma_0/J$. For each hardening rate, this can be represented by a family of curves, as shown graphically in Fig. 8 for $n = 13$. The results are also summarized by equations (8) and (9).

$$\left(\frac{\sigma_{\theta\theta}}{\sigma_0}\right)_{(r,T)} = \left(\frac{\sigma_{\theta\theta}}{\sigma_0}\right)_{(r,T=0)} + 0.64\left(\frac{T}{\sigma_0}\right) - 0.4\left(\frac{T}{\sigma_0}\right)^2 \quad n = 13, \frac{T}{\sigma_0} \le 0 \tag{8}$$

$$\left(\frac{\sigma_{\theta\theta}}{\sigma_0}\right)_{(r,T)} = \left(\frac{\sigma_{\theta\theta}}{\sigma_0}\right)_{(r,T=0)} + 0.6\left(\frac{T}{\sigma_0}\right) - 0.75\left(\frac{T}{\sigma_0}\right)^2 \quad n = \infty, \frac{T}{\sigma_0} \le 0. \tag{9}$$

The original HRR analyses were asymptotic expansions of the crack-tip field for a nonlinear elastic solid, in which solutions were restricted to the first dominant term of the series. More recently, Li and Wang (1986) have sought two-term asymptotic expansions of the form:

$$\sigma_{ij} = K_{s1} r^{s1} f_{ij}(\theta, n) + K_{s2} r^{s2} g_{ij}(\theta, n) \tag{10}$$

where the first term is identified with the HRR field, and numerical calculated values have been given for the exponent of the second term, for hardening rates $n = 3$, and $n = 10$. The present results indicate that the second term in the elastic-plastic expansion is a function of T, which is the second term in the elastic expansion. The strength of this term is so much weaker than that of the first term in the series that the second term is justified as being considered to be independent of $r\sigma_0/J$, as summarized in equations (8) and (9).

In contrast, the effect of positive T stresses is to raise the stress levels slightly until they approach those of the HRR field more closely. However, in this case the family of curves shown in Fig. 8 are not parallel but rotate slightly indicating the presence of an r dependence. Nevertheless, the effect of positive T stresses is so much weaker than negative T stresses that, for all practical purposes, fields with zero or positive T stresses can be regarded as being close to the HRR field and characterized by J alone. To examine the J dominance criteria suggested by Shih and German (1981), the stresses may be compared with the HRR field for $n = 13$, with the criterion that J dominated flow fields have stresses within 90 percent of the HRR value at a distance $r = 2J/\sigma_0$. On this basis J

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

dominance is maintained for T stresses greater than $-0.2\sigma_0$ for $n = 13$.

Using biaxiality values given in the literature, the J dominance criterion may now be applied to a wide range of geometries. Four geometries are considered as examples: center-cracked panels (CCP); single-edge notched bend bars (SENB); force-loaded single-edge notched tension bars (SENT); and double-edge cracked bars in tension (DECT). In general, deeply cracked edge-cracked bars exhibit positive biaxialities, while the same geometries exhibit negative B values for short cracks. The transition from shallow to deeply cracked geometries occurs for edge-cracked bars at $a/W = 0.3$ in bending and 0.55 in tension and at 0.9 for the double edge-cracked geometry. These transitions also match the transitions in the form of the corresponding slip-line field fields, indicating an interesting correlation between the plastic and elastic solutions, which is discussed in more detail elsewhere (Al-Ani and Hancock, 1989). The center-cracked panel, however, exhibits strongly negative B values for all (a/W) ratios.

The choice of a J dominance criterion is somewhat subjective. However, if that proposed by Shih and German (1981) is adopted, the size limitations for J dominated flow are given in Fig. 16. These requirements are exceptionally severe and lead to the conclusion that central cracked panels lose J dominance within the requirements of L.E.F.M. as codified in both the British and American standards (A.S.T.M., 1970). If comparison were to be made with the $B = 0$ field the results would be less severe, but would exhibit the same trends. Specimens with positive T stresses maintain J dominance until the crack-tip opening becomes a significant fraction of the ligament when J dominance is lost at $25J/\sigma_0$, due to large geometry change effects.

Conclusions

Elastic stress fields can be characterized by the stress intensity factor K and the biaxiality parameter B. In large-scale yielding elastic-plastic fields can also be characterized by J and T into full plasticity. The two-term characterization of elastic-plastic fields corresponds to the introduction of a second term in the asymptotic crack-tip expansions, first introduced by Hutchinson (1968a,b) and Rice and Rosengren (1968). The second term is so weakly dependent on positive values of the biaxiality parameter B that the stress fields can, for all practical purposes, be described by the HRR field and characterized by J alone. This observation brings together all the geometries which show J dominance. Similarly, the geometries which lose J dominance feature negative T stresses, and are characterised by two parameters, J and T. These observations now unify the problems of J dominance in different geometries until large geometry change effects associated with crack-tip blunting become significant at $25J/\sigma_0$. It is no longer necessary to have different criteria for different configurations and loading conditions. J dominance is now simply a special case of two-parameters characterization of crack-tip fields, for zero or positive T stresses, such that the effect of the second term in the nonlinear asymptotic expansion of the crack-tip field is negligible.

Acknowledgments

C. Betegón is pleased to be able to acknowledge the support of a grant from F.I.C.Y.T., and thanks are due to Hibbitt Karlson and Sorensen, Inc. for access to ABAQUS. Both authors take pleasure in being able to acknowlege the value of discussions with Prof. D. M. Parks, C. F. Shih, R. McMeeking, and J. Sumpter.

References

Al-Ani, A. A., and Hancock, J. W., 1991, "J-Dominance of Short Cracks in Tension and Bending," *Journal of the Mechanics and Physics of Solids*, Vol. 39, pp. 23–43.

A.S.T.M., 1970, *Annual Book of A.S.T.M. Standards*, A.S.T.M. Society for Testing and Material, Philadelphia, Pa., pt. 31, pp. 911.

Bilby, B. A., Cardew, G. E., Goldthorpe, M. R., and Howard, I. C., 1986, "A Finite element Investigation of the Effect of Specimen Geometry on the Fields of Stress and Strain at the Tip of Stationary Cracks," *Size Effects in Fracture*, I. Mech. E., London, pp. 37–46.

Bilby, B. A., Cardew, G. E., Goldthorpe, M. R., and Howard, I. C., 1987, "Fifth Report on the Stability of Cracks in Tough Materials: A Finite Element Investigation on the Effect of Specimen Geometry on Crack Tip Fields in Hardening Materials," Department of Mech. Eng, University of Sheffield, Sheffield, U.K.

Cardew, G. E., Goldthorpe, M. R., Howard, I. C., and Kfouri, A. P., 1984, "On the Elastic T-Term," *Fundamentals of Deformation and Fracture* (Eshelby Memorial Symposium), Cambridge University Press, pp. 465–476.

Hibbitt, Karlsson and Sorensen, 1984a, *ABAQUS Theory Manual*, Providence, R. I.

Hibbitt, Karlsson and Sorensen, 1984b, *ABAQUS Users Manual*, Providence, R. I.

Hutchinson, J. W., 1968a, "Singular Behavior at the End of a Tensile Crack in a Hardening Material," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 13–31.

Hutchinson, J. W., 1968b, "Plastic Stress and Strain Fields at a Crack Tip," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 337–347.

James-Watt, 1989, "T Stresses in Fracture Mechanics," *Internal Report*, Department of Mechanical Engineering University of Glasgow.

Kfouri, A. P., 1986, "Some Evaluations of the Elastic T-term using Eshelby's Method," *International Journal of Fracture*, Vol. 30, pp. 301-315.

Larsson, S. G., and Carlsson, A. J., 1973, "Influence of non-singular Stress Terms and Specimen Geometry on Small-Scale Yielding at Crack Tips in Elastic-Plastic Material," *Journal of the Mechanics and Physics of Solids*, Vol. 21, pp. 263–278.

Leevers, P. S., and Radon, J. C., 1983, "Inherent Stress Biaxiality in various Fracture Specimen Geometries," *International Journal of Fracture*, Vol. 19, pp. 311–325.

Li F. Z., Shih C. F., and Needleman A., 1985, "A Comparison of Methods for Calculating Energy Release Rate," *Eng. Fract. Mech.* Vol. 21, pp. 405–421.

Li, Y., and Wang, Z., 1986, "High-Order Asymptotic field of Tensile Plane-Strain Nonlinear crack Problems," *Scientia Sinica* (Series A), Vol. 29, pp. 942–955.

McClintock F. A., 1971, "Plasticity Aspects of Fracture," *Fracture*, Vol. 3, H. Liebowitz, eds., Academic Press, London, pp. 47–225.

McMeeking, R. M., 1977, "Finite Deformation Analysis of Crack-Tip Opening in Elastic-Plastic Materials and Implications for Fracture," *Journal of the Mechanics and Physics of Solids*, Vol. 25, pp. 375–381.

McMeeking, R. M., and Parks, D. M., 1979, "On Criteria for J-Dominance of Crack-Tip Fields in Large-Scale Yielding," *Elastic-Plastic Fracture Mechanics*, ASTM-STP 668, ASTM, Philadelphia, Pa., pp. 175–194.

Parks, D. M., 1977, The Virtual Crack Extension Method for Non-Linear Material Behavior," *Computer Methods in Applied Mechanics and Engineering*, Vol. 12, pp. 353–364.

Parks, D. M., and Wang, Y. Y., 1988, "Elastic-Plastic Analysis of Part-Through Surface Cracks," *Proc. ASME Symp. on 3-Dimensional Effects in Fracture*, ASME, New York.

Rice, J. R., 1968, "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 379–386.

Rice, J. R., 1974, "Limitations to the Small Scale Yielding Approximation for Crack Tip Plasticity," *Journal of the Mechanics and Physics of Solids*, Vol. 22, pp. 17–26.

Rice, J. R., and Rosengren, G. F., 1968, "Plane Strain Deformation Near a Crack Tip in a Power Law Hardening Material," *Journal of the Mechanics and Physics of Solids*, Vol. 16, pp. 1–12.

Shih, C. F., 1983, "Tables of Hutchinson-Rice-Rosengren Singular Field Quantities," Brown University Report, MRL E-147.

Shih, C. F., and German, M. D., 1981, "Requirements for a one Parameter Characterization of Crack Tip Fields by the HRR Singularity," *International Journal of Fracture*, Vol. 17, pp. 27–43.

Tracey, D., 1976, "Finite Elements Solutions for Crack-Tip Behavior in Small-Scale Yielding," *ASME Journal of Engineering Materials and Technology*, Vol. 98, pp. 146–151.

Williams, M. L., 1957, "On the Stress Distribution at the Base of a Stationary Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, pp. 111–114.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Near-Tip Fields for Quasi-Static Crack Growth Along a Ductile-Brittle Interface[1]

W. J. Drugan
Department of Engineering Mechanics,
University of Wisconsin,
Madison, WI 53706
Assoc. Mem. ASME

As an analytical first study of the mechanics of ductile-brittle interfacial crack growth, I derive the stress and deformation fields near the tip of a crack that is growing quasi-statically along an interface joining material phases idealized as being rigid on one side and homogeneous, isotropic, elastic-ideally plastic on the other. Both the cases of antiplane strain and plane strain are treated, since it is some combination of these that pertains near an interface crack under general three-dimensional conditions. In the antiplane strain case, a family of solutions for the growing crack fields is found covering all admissible crack line shear stress ratios except for the case when the shearing is parallel to the crack; the Mode III homogeneous material solution is shown to be a limiting member of this family. Two distinct solution families are shown to exist for the growing crack fields in the plane strain case. One of these contains the Mode I homogeneous material solution and is characterized by large crack line stress triaxialities (up to 39 percent higher than that in homogeneous material solutions), while the other family has a limiting member that closely resembles (but is not identical to) the Mode II homogeneous material solution and is characterized by small and even negative crack line stress triaxialities. These solutions, which appear to constitute all possible that involve traction-free crack faces, are shown to cover large portions, but not the entire admissible range, of crack line near-tip shear/normal stress ratios ("mixities").

1 Introduction

Interfaces between ductile and brittle materials occur in a variety of advanced materials and in other technological situations. For example, ductile materials (e.g., metals, polymers) are often strengthened by the presence of strong brittle particles or fibers, whereas strong brittle materials, such as ceramics, are toughened against fracture by the dispersion of ductile (e.g., metal) particles. Many electronic devices involve a thin metal film bonded to a brittle substrate; they also utilize solder joints. In all of these cases, debonding along the ductile-brittle interface can occur; this may be undesirable or desirable depending on the situation (for example, it may produce increased toughness in brittle materials). Whatever the circumstance, a clear understanding of the mechanics of the debonding process is pertinent to maximizing structural efficiency and/or component reliability.

In the present work I attempt to elucidate fundamental mechanics aspects of this debonding process for situations when the size scale is sufficiently large to warrant a continuum mechanics approach to determine the stress and deformation fields in the vicinity of a quasi-statically extending brittle/ductile interface debond (i.e., interface crack). The constitutive assumptions employed are physically sensible yet sufficiently idealized to facilitate closed-form analytical results: The brittle material is modeled as being rigid, while the ductile material is taken to be homogeneous, incompressible, isotropic elastic-ideally plastic Prandtl-Reuss-Mises material. Recognizing that the near-tip state for an interface crack inside a loaded structural component experiencing general three-dimensional conditions is in general some combination of plane and antiplane strain deformations, since it is the plane and antiplane components of the strain tensor that become infinite at a three-dimensional crack front but not (generally) the component of strain parallel to the crack front (Rice, 1982), both of these cases are analyzed here.

Most previous mechanics analyses of interface cracks have consisted of linear elastic solutions for crack tip fields: Williams (1959), Erdogan (1963, 1965), Sih and Rice (1964), Rice and Sih (1965), and England (1965) considered cracks between dissimilar isotropic linear elastic materials, and Willis (1971) and Clements (1971) treated cracks between dissimilar anisotropic linear elastic materials. These early works have been clarified and extended in several more recent contributions such as those of Park and Earmme (1986), Hutchinson et al. (1987), and Rice (1988). It has only been very recently that plastic deformation in at least one of the interface materials has been treated; analyses for stationary

[1] Presented at the Third Joint ASCE-ASME Mechanics Conference, University of California, San Diego, Calif., July 9–12, 1989.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 2, 1989; final revision, Feb. 13, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

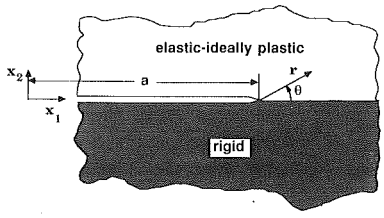


Fig. 1 Interface crack geometry. The lower material is assumed rigid and immobile, while the upper material is homogeneous, isotropic, incompressible, elastic-ideally plastic. Cartesian coordinates x_i are fixed in the body; polar coordinates r, θ are centered at the crack tip and move with it through the material as the crack grows.

cracks include those of Shih and Asaro (1988, 1989a,b), Shih et al. (1990a,b), Parks and Zywicz (1989), and Zywicz and Parks (1990). To my knowledge, the only analyses of elastic-plastic interfacial crack *growth* have been numerical studies, e.g., those of Needleman (1990a,b) and Varias et al. (1990), which employ a cohesive theory to model the interface between rigid or elastic and viscoplastic materials.

2 Formulation

The geometry of both cases to be treated is illustrated in Fig. 1. A crack grows quasi-statically along an elastic-plastic/rigid interface which is taken to be locally planar. The Cartesian coordinate system **x** is fixed in the body such that x_1 points in the crack growth direction and x_3 is parallel to the (straight) crack front. Crack length is a, so that $\dot{a}$ is crack growth velocity. A polar coordinate system r, θ lies in the x_1-x_2 plane, is centered at the crack tip and moves with it through the material as the crack grows, θ being measured counterclockwise from the interface ahead of the crack. Thus, the nonrigid material lies in $0 \le \theta \le \pi$, so this will be the domain of our solutions.

The governing equations are those that pertain to the quasi-static deformation of fully incompressible, nonhardening Prandtl-Reuss-Mises material within a small-displacement-gradient formulation, i.e., neglecting the effects of deformation-produced geometry changes. Equilibrium requires the stress tensor to be symmetric and to satisfy

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{b} = \mathbf{0}, \tag{1}$$

where ∇ is the usual vectorial differentiation operator so that $\nabla\cdot$ indicates divergence, $\boldsymbol{\sigma}$ is the stress tensor, and **b** is the body force vector per unit volume. Geometrical compatibility relates the rate-of-deformation tensor $\mathbf{D}$ ($= \dot{\boldsymbol{\epsilon}}$ for $|\nabla \mathbf{u}| \ll 1$, where $\dot{\boldsymbol{\epsilon}}$ is time rate of the infinitesimal strain tensor and **u** is the displacement vector) to the material velocity vector **v** as

$$\mathbf{D} = [\nabla \mathbf{v} + (\nabla \mathbf{v})^T]/2, \tag{2}$$

where a superscript T denotes transpose. Elastic-plastic material response is assumed to be described by the Prandtl-Reuss flow rule

$$\mathbf{D} = \mathbf{D}^e + \mathbf{D}^p = \frac{3}{2E}\,\dot{\mathbf{s}} + \dot{\Lambda}\mathbf{s}, \tag{3}$$

where superscripts e and p denote elastic and plastic parts, respectively; E is the Young modulus; $\mathbf{s} = \boldsymbol{\sigma} - \text{trace}\,(\boldsymbol{\sigma})\mathbf{I}/3$ is the deviatoric stress tensor with **I** denoting the identity tensor; $\dot{\Lambda} \ge 0$ is an undetermined parameter; and a superposed dot denotes time rate at a fixed material point. Material is at yield when the Huber-Mises yield condition is met:

$$\sigma_e \equiv [(3/2)\mathbf{s}:\mathbf{s}]^{1/2} = \sqrt{3}k, \tag{4}$$

where σ_e is the effective stress and k is the (constant) yield stress in shear. Thus, in regions of ongoing plastic deformation, equations (3) and (4) must be satisfied with $\dot{\Lambda} > 0$; in material experiencing instantaneously elastic response, whether or not it has previously deformed plastically, (3) governs with $\dot{\Lambda} \equiv 0$ and (4) applies with $\le$ replacing $=$.

Rice (1982) argued, and Drugan (1985) proved, that these governing equations become analytically tractable in the limit as $r \to 0$ even for general unsteady crack growth. Specifically, if **b** is assumed bounded everywhere and mild, plausible assumptions are made concerning the existence of certain quantities involving stress derivatives as $r \to 0$, Drugan (1985) proved that equations (1) and (3) simplify to the following asymptotic forms, specialized here to the present problems in which all field quantities are assumed independent of x_3:

$$\partial\sigma_{r\theta}/\partial\theta + \sigma_{rr} - \sigma_{\theta\theta} + o\{[\ln(R/r)]^{-1}\} = 0 \ \text{ as } r \to 0 \tag{5a}$$

$$\partial\sigma_{\theta\theta}/\partial\theta + 2\sigma_{r\theta} + o\{[\ln(R/r)]^{-1}\} = 0 \ \text{ as } r \to 0 \tag{5b}$$

$$\partial\sigma_{\theta 3}/\partial\theta + \sigma_{r3} + o\{[\ln(R/r)]^{-1}\} = 0 \ \text{ as } r \to 0 \tag{5c}$$

$$D_{ij} = (3/2)(\dot{a}/E)\,(\sin\theta/r)\partial s_{ij}/\partial\theta + \dot{\Lambda}s_{ij} + o\{[r\ln(R/r)]^{-1}\} \quad \text{as } r \to 0, \tag{6}$$

respectively, where R is an undetermined parameter having length dimensions. Here and in the sequel, Roman indices i,j, k,l have range 1, 2, 3, while Greek indices α, β, γ have range 1, 2 only; both sets of indices represent components of tensors with respect to the Cartesian coordinate system of Fig. 1 and obey the usual summation convention. The indices r, θ represent tensor components with respect to the polar coordinate system of Fig. 1 and do *not* follow a summation convention.

As will be reviewed explicitly in the following sections, in each of the cases of antiplane strain and plane strain, the asymptotic ($r \to 0$) form of this governing equation set was shown by Rice (1982) to admit three types of solution—that is, three types of near-tip angular sector (two types of plastically deforming sector and an instantaneously elastic sector). Thus, the construction of complete near-tip solutions for the growing crack stress and deformation fields involves assembling sectors of these permissible types in such a way that plastic dissipation is non-negative everywhere, the yield condition is not violated in instantaneously elastic sectors, and intersector continuity conditions, and the problem's boundary conditions, are satisfied.

Since the crack is growing, intersector boundaries are propagating quasi-statically through the material. A general study of the restrictions on possible jumps in components of stress, strain, and material velocity across such a propagating surface in a broad class of elastic-plastic solids, inclusive of the model employed here, was conducted by Drugan and Rice (1984) and improved by Drugan (1986). Specialized to the present constitutive model and deformation modes, their results provide the following important requirements across intersector boundaries: (i) *all* components of the stress tensor must be continuous; (ii) only the material velocity components *parallel* to the moving surface, and their associated shear strain components, may experience jumps, *provided* the following conditions are met:

$$s^{\Sigma}_{tt} = s^{\Sigma}_{nn} = \sigma^{\Sigma}_{t3} = 0 \tag{7}$$

$$[\![v_\delta]\!] = -2V^*[\![\Lambda]\!]\sigma^{\Sigma}_{n\delta} \tag{8}$$

$$[\![\epsilon^p_{n\delta}]\!] = [\![\Lambda]\!]\sigma^{\Sigma}_{n\delta}. \tag{9}$$

Here, a superscript Σ indicates that the (continuous) stress component is to be evaluated on the potential discontinuity surface; subscripts n and t refer to normal (in the propagation direction) and tangential (in the x_1, x_2 plane) directions to the moving surface whose normal velocity is V^*, and subscript δ stands for t and 3; $[\![\ ]\!]$ denotes the jump in a quantity across the surface (value just ahead minus value just behind); and $[\![\Lambda]\!] \le 0$.

The boundary conditions to be imposed are as follows: Without loss of generality, we assume the rigid

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

material in Fig. 1 to remain motionless. Thus, the boundary condition along $\theta = 0$ is the requirement that the displacement vector **u**, and hence the material velocity vector **v**, vanish. The boundary condition along the crack face $\theta = \pi$ is that the traction vector $\mathbf{T} = \mathbf{0}$ (our solutions will verify that no material interpenetration is predicted along the crack flanks).

3 Asymptotic Growing Crack Fields in Antiplane Strain

In the simple case of antiplane strain deformations, the only nonvanishing component of velocity is $v_3 = v_3(r,\theta)$. From this, equations (1)–(3) show that the only nonvanishing components of σ, **D**, and ϵ are the 31 (=13) and 32 (=23) components, which must also be independent of x_3. Thus, near a growing crack tip (i.e., for $r \to 0$), the requirements of equilibrium and the flow rule reduce to (5*c*) and the $\alpha 3$ components of (6), respectively; compatibility is ensured by enforcement of the $\alpha 3$ components of (2); and the yield condition (4) simplifies to:

$$\sigma_{13}^2 + \sigma_{23}^2 = k^2. \tag{10}$$

To leading order in r (i.e., neglecting higher-order terms as $r \to 0$ in equations (5*c*) and (6)), Rice (1982) showed that the equation set summarized in the preceding sentence admits only three types of solution—that is, three types of near-tip angular sector. These are summarized below (neglecting arbitrary rigid-body motion) with the sign of $\sigma_{\theta 3}$ in the "centered fan" plastic sector chosen positive, as it will be in our solutions.

(*i*) *"Centered Fan" Plastically Deforming Sector.*

$$\sigma_{r3} = 0, \quad \sigma_{\theta 3} = k \quad [\Rightarrow \sigma_{13} = -k \sin\theta, \ \sigma_{23} = k\cos\theta] \tag{11}$$

$$v_3 = (k/G)\dot{a}\sin\theta \ln(R/r) + f(\theta,t) \tag{12}$$

$$\dot{\Lambda} = (1/2G)\cos\theta(\dot{a}/r)\ln(R/r); \tag{13}$$

(*ii*) *"Constant Stress" Plastically Deforming Sector.*

$$\sigma_{\alpha 3} = Q_{\alpha 3} k \tag{14}$$

$$v_3 = -(k/G)B^*\dot{a}[\ln(R/r) - \ln|Q_{13}\cos\theta + Q_{23}\sin\theta|] \tag{15}$$

$$\dot{\Lambda} = (B^*/2G)(\dot{a}/r)/(Q_{13}\cos\theta + Q_{23}\sin\theta); \tag{16}$$

(*iii*) *Elastically Deforming Sector.*

$$\sigma_{13} = k[C + B \ln|\sin\theta|], \quad \sigma_{23} = k[A + B\theta] \tag{17}$$

$$v_3 = -(k/G)B\dot{a}\ln(R/r). \tag{18}$$

In the preceding equations, A, B, B^*, C, and $Q_{\alpha 3}$ are undetermined dimensionless constants (except that $Q_{\alpha 3}$ must be such that equation (14) satisfies (10)); G ($=E/3$) is the elastic shear modulus; $f(\theta, t)$ is an undetermined function of integration; and t is time or some monotonically increasing parameter. Equations (15) and (16) were derived by assuming that the velocity in a "constant stress" plastically deforming sector is logarithmically singular, as it must be in the other sector types.

Now, as noted earlier, solutions to the present antiplane strain problem of determining the near-tip growing interface crack stress and deformation fields in the ductile material are constructed by assembling angular sectors chosen from the three permissible types summarized previously so that plastic dissipation is non-negative everywhere, the yield condition is not violated in instantaneously elastic sectors, and the following intersector jump conditions and problem boundary conditions are satisfied (specializing the forms given in Section 2):

$$[\![\sigma_{\alpha 3}]\!] = 0 \tag{19}$$

$$\sigma_{r3}^{\Sigma} = 0 \text{ and } [\![v_3]\!] = 2V^*[\![\Lambda]\!]\sigma_{\theta 3}^{\Sigma}; \text{ otherwise } [\![v_3]\!] = 0 \tag{20}$$

$$v_3 = 0 \text{ on } \theta = 0 \tag{21}$$

$$\sigma_{23} = 0 \text{ on } \theta = \pi. \tag{22}$$

Since the near-tip boundary conditions do not specify the stress state on the line ahead of the crack, we expect that different near-tip solutions will exist for different ratios of the far-field applied stresses σ_{23}/σ_{13}. In a manner similar to that of Shih (1974) who treated plane strain mixed-mode stationary crack tip fields, I define a *near-tip* "mixity parameter" M_a (the subscript *a* denoting antiplane strain) as

$$M_a \equiv \frac{2}{\pi}\arctan\left[\lim_{r\to 0}\frac{\sigma_{23}(r,\theta=0)}{-\sigma_{13}(r,\theta=0)}\right]. \tag{23}$$

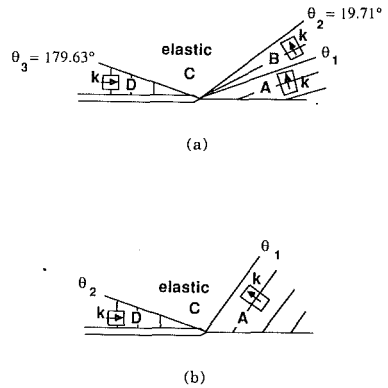


Fig. 2 Near-tip solution configurations in the ductile material for antiplane strain interfacial crack growth; stress characteristics are shown in plastic sectors. (*a*) applies for $1 \geq M_a \geq 0.7810$ ($0 \leq \theta_1 \leq 19.71$ deg), while (*b*) applies for $0.7810 \geq M_a > 0$ ($19.71 \text{ deg} \leq \theta_1 < 90$ deg).

I shall now show that near-tip growing crack solutions can be found for all values of M_a in the range $0 < M_a \leq 1$; it appears that full-field (numerical) solutions are necessary to determine which value of M_a corresponds to a given ratio of far-field applied stress.

To begin, we observe that a subyield sector cannot border $\theta = 0$. This is due to boundary condition (21) which requires $B = 0$ in equation (18); this shows via (17) that the Cartesian stress components in such a potential subyield sector must be constants, but since this sector must border a sector at yield for a nontrivial ($v_3 \neq 0$) solution, and due to (19), we conclude that these constant stresses must satisfy yield.

Thus, the sector bordering $\theta = 0$ must be of the centered fan or constant stress plastic type (observing that the latter with $B^* = 0$ is indistinguishable from an elastic sector that is at yield with $B = 0$). More specifically, a centered fan sector may border $\theta = 0$ only for the case $M_a = 1$ [$\sigma_{13}(r, \theta = 0) = 0^-$], as is evident from equation (11); this corresponds to the same stress state on the line ahead of the crack as the Chitaley and McClintock (1971) solution for pure Mode III quasi-static crack growth in *homogeneous* elastic-ideally plastic material. Indeed, it is easy to verify that their solution solves the present problem in the ductile material for this $M_a = 1$ case, involving as it does a centered fan plastic sector bordering $\theta = 0$, which via (12) (taking $f(0,t) = 0$) is seen to satisfy boundary condition (21) (and all other requisite conditions, which are the same as those in Chitaley and McClintock's (1971) problem). This solution will be reviewed shortly.

For $M_a < 1$, permissible solutions must have a constant stress plastic sector bordering $\theta = 0$. Two near-tip solution configurations with this property have been found that satisfy all necessary conditions outlined above; they are illustrated in Fig. 2. Figure 2(*a*) illustrates the solution configuration when M_a is in the range $0.78098640 \leq M_a \leq 1$. This configuration involves a constant stress plastic sector (A) directly ahead of the growing crack, followed by a centered fan plastic sector (B), an elastically deforming sector (C), and finally another constant stress plastic sector (D). Application of the jump and boundary conditions (19)–(22), and also the requirement that v_3 be continuous across θ_2 (which can be proved when θ_2 is a centered fan—elastic sector boundary by enforcing the requirements of non-negative plastic work produced by a

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Values of the parameters appearing in equations (30)–(33) for the antiplane strain solutions corresponding to Fig. 2(*b*)

θ_1 (deg)	θ_2 (deg)	Q^A_{13}	Q^A_{23}	A	B	C	M_a
19.711224	179.63341	−0.33727969	0.94140449	1.0574374	−0.33727969	−0.70384967	0.78098640
30.0	179.68964	−0.5	0.86602540	1.0395896	−0.33148319	−0.72976664	0.66666667
40.0	179.80336	−0.64278761	0.76604444	0.98522214	−0.31394892	−0.78153449	0.55555556
50.0	179.91353	−0.76604444	0.64278761	0.89017816	−0.28348869	−0.84159846	0.44444444
60.0	179.97997	−0.86602540	0.5	0.75004173	−0.23877226	−0.90037065	0.33333333
70.0	179.99899	−0.93969262	0.34202014	0.55967132	−0.17814991	−0.95077398	0.22222222
80.0	180.0⁻	−0.98480775	0.17364818	0.31256672	−0.09949308	−0.98633087	0.11111111
90.0	180.0	−1.0	0.0	0.0	0.0	−1.0	0.0

hypothesized velocity jump and of nonviolation of yield by the stress state in the elastically deforming sector in a manner completely analogous to the plane strain version of Drugan and Chen (1989, Appendix A)) to the intrasector solutions (11)–(18) results in the following equations for θ_2 and θ_3:

$$\sin\theta_2 \ln\left|\frac{\sin(\theta_2+\cot\theta_2)}{\sin\theta_2}\right| + \sin\theta_2 + 1 = 0 \tag{24}$$

$$\theta_3 = \theta_2 + \cot\theta_2. \tag{25}$$

Only one physically meaningful solution set exists to these equations; this is given below, together with all other pertinent quantities relating to the solution family of Fig. 2(*a*):

$$\theta_1 = (1-M_a)\pi/2,\ \theta_2 = 0.34402576 (= 19.711224 \text{ deg}),$$
$$\theta_3 = 3.1351945 (= 179.63341 \text{ deg}) \tag{26}$$

$$Q^A_{13} = -\sin\theta_1,\ Q^A_{23} = \cos\theta_1,\ B^{*A} = 0 \tag{27}$$

$$\left.\begin{aligned} A &= \theta_3\sin\theta_2 = 1.0574374, \quad B = -\sin\theta_2 = -0.33727969 \\ C &= 1 + \sin\theta_2 \ln|\sin\theta_3| = -0.70384967 \end{aligned}\right\} \tag{28}$$

$$Q^D_{13} = 1, \quad Q^D_{23} = 0, \quad B^{*D} = B. \tag{29}$$

Thus, the near-tip stress and velocity fields for antiplane strain interfacial crack growth when M_a is in the range $0.78098640 \le M_a \le 1$ are given in accordance with Figure 2(*a*) by equations (11)–(18) with the parameters appearing in those expressions being specified by (26)–(29), in which superscripts A and D denote parameters valid in Sectors A and D, respectively.

Several important observations follow from these equations. First, they give the Chitaley and McClintock (1971) solution when $M_a = 1$. (Note that Chitaley and McClintock's (1971) values for what here are called θ_2 and θ_3 are correct only to three significant figures.) Second, equation (26) shows that θ_2 and θ_3 are independent of M_a, and hence (26)–(29) show that for M_a in this range, the near-tip fields change only in the region $0 \le \theta \le \theta_1$. Third, except for the $M_a = 1$ case, these solutions involve zero velocity (at least to first order in r) in the leading constant stress plastic sector and hence a discontinuity in v_3 across θ_1, which is permissible since it satisfies (20) and produces positive plastic work.

For $M_a = 0.78098640$, equations (26) show $\theta_1 = \theta_2$, so that the centered fan has vanished. This is a hint that the resulting simpler configuration, shown in Fig. 2(*b*), will give the solution family for the remaining range of $0 < M_a \le 0.78098640$. This is indeed the case. This configuration has a leading constant stress plastic sector (A) bordered by an elastically deforming sector (C), which is followed by another constant stress plastic sector (D). Application of conditions (19)–(22) to (14)–(18) for this configuration leads to the following results:

$$\theta_1 = (1-M_a)\pi/2 \tag{30}$$

$$\cos\theta_1 \ln|\sin\theta_2/\sin\theta_1| + (\theta_2-\theta_1)(1+\sin\theta_1) = 0 \tag{31}$$

$$Q^A_{13} = -\sin\theta_1,\ Q^A_{23} = \cos\theta_1,\ B^{*A} = 0 \tag{32}$$

$$B = \cos\theta_1/(\theta_1-\theta_2),\ A = -B\theta_2,\ C = 1 - B\ln|\sin\theta_2| \tag{33}$$

$$Q^D_{13} = 1,\ Q^D_{23} = 0,\ B^{*D} = B. \tag{34}$$

Thus, the near-tip stress and velocity fields in the ductile material for antiplane strain interfacial crack growth when M_a is in the range $0 < M_a \le 0.78098640$ are given in accordance with Fig. 2(*b*) by equations (14)–(18) with parameters specified by (30)–(34), in which superscripts A and D denote parameters valid in Sectors A and D, respectively. Observe that (31) is an implicit equation for θ_2 for a given θ_1 (i.e., a given M_a). This solution family also involves zero velocity in the leading constant stress plastic sector and hence a jump in v_3 across θ_1, which is permitted by (20) and which produces positive plastic work. (Solutions having the configuration of Fig. 2(*b*) cannot be found unless v_3 jumps across θ_1; (20) and (10) thus require $|\sigma_{\theta 3}(\theta_1)| = k$, i.e., θ_1 must be a stress characteristic.) Values of the parameters given in (30)–(33) are displayed in Table 1 for representative values of M_a in the range of this solution family's validity. It is interesting to note that solutions to (31) can also be found for $M_a > 0.78098640$ ($\theta_1 < 19.711224$ deg), but such solutions are inadmissible because the yield condition is violated in Sector C. Observe that a growing crack solution with a nonzero crack face sliding displacement does not exist for $M_a = 0$, since $B^{*D} = 0$ in that case.

Finally, it is easy to show that an admissible family of growing crack near-tip solutions can also be found for the range $-1 \le M_a < 0$: this family involves identical near-tip configurations to those described above; the differences are, briefly, that the stress and velocity components all change sign.

4 Asymptotic Growing Crack Fields in Plane Strain

4.1 Near-Tip Solution Sectors; Continuity and Boundary Conditions. For plane strain deformations, the nonvanishing components of material velocity are $v_1 = v_1(r, \theta)$ and $v_2 = v_2(r,\theta)$. From these, equations (1)–(3) show that the only nonvanishing components of σ, $\mathbf{D}$, and ϵ are the 11, 12 (=21), and 22 components [and $\sigma_{33} = (\sigma_{11}+\sigma_{22})/2$], which must also be independent of x_3. Thus, in the ductile material near a growing crack tip (i.e., for $r \to 0$), the requirements of equilibrium and the flow rule reduce to (5*a*,*b*) and the $\alpha\beta$ components of (6), respectively; compatibility is ensured by enforcement of the $\alpha\beta$ components of (2); and the yield condition (4) simplifies to:

$$(\sigma_{11}-\sigma_{22})^2/4 + \sigma_{12}^2 = k^2. \tag{35}$$

To leading order in r (i.e., neglecting the higher-order terms as $r \to 0$ in equations (5*a*,*b*) and (6)), Rice (1982) showed that the equation set summarized in the preceding sentence admits only three types of solution—that is, three types of near-tip angular sector. These are summarized as follows (neglecting arbitrary rigid-body motion) with the sign of $\sigma_{r\theta}$ in the centered fan plastic sector chosen positive, as it will be in our solutions.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(*i*) *"Centered Fan" Plastically Deforming Sector.*

$$\sigma_{r\theta} = k, \quad \sigma_{rr} = \sigma_{\theta\theta} = \sigma_{33} = (C - 2\theta)k \tag{36}$$

$$v_r = 3(k/E)\dot{a}\sin\theta \ln(R/r) + \partial f(\theta,t)/\partial\theta \tag{37}$$

$$v_\theta = 3(k/E)\dot{a}(\cos\theta - V)\ln[R/(er)] - f(\theta,t) \tag{38}$$

$$\dot{\Lambda} = (3/2)(V/E)(\dot{a}/r)\ln(R/r); \tag{39}$$

(*ii*) *"Constant Stress" Plastically Deforming Sector.*

$$\sigma_{\alpha\beta} = P_{\alpha\beta}k, \quad \sigma_{33} = (P_{11} + P_{22})k/2 \tag{40}$$

$$v_1 = -\dot{a}(k/E)B_1 \ln|R/(r\sin\theta)| - \dot{a}\int[F_{11}(\theta)/\sin\theta]d\theta \tag{41}$$

$$v_2 = -\dot{a}(k/E)B_2 \ln|R/(r\cos\theta)| + \dot{a}\int[F_{22}(\theta)/\cos\theta]d\theta \tag{42}$$

$$\dot{\Lambda} = \frac{(2/E)(B_1\cos\theta + B_2\sin\theta)(\dot{a}/r)}{(P_{11} - P_{22})\cos 2\theta + 2P_{12}\sin 2\theta} \tag{43}$$

$$F_{\alpha\beta}(\theta) = (kr\dot{\Lambda}/\dot{a})(P_{\alpha\beta} - \delta_{\alpha\beta}P_{\gamma\gamma}/2); \tag{44}$$

(*iii*) *Elastically Deforming Sector.*

$$\sigma_{11} = (k/3)[4A_1 \ln|\sin\theta| + A_1\cos 2\theta + A_2(2\theta + \sin 2\theta) + C_{11}] \tag{45}$$

$$\sigma_{22} = (k/3)[-A_1\cos 2\theta + A_2(2\theta - \sin 2\theta) + C_{22}] \tag{46}$$

$$\sigma_{12} = (k/3)[A_1(2\theta + \sin 2\theta) - A_2\cos 2\theta + C_{12}] \tag{47}$$

$$\sigma_{33} = (\sigma_{11} + \sigma_{22})/2 \tag{48}$$

$$v_1 = -\dot{a}(k/E)[A_1\ln(R/r) - A_1\sin^2\theta + A_2\sin\theta\cos\theta] \tag{49}$$

$$v_2 = -\dot{a}(k/E)[A_2\ln(R/r) + A_1\sin\theta\cos\theta + A_2\sin^2\theta]. \tag{50}$$

In the preceding equations, A_α, B_α, C, $C_{\alpha\beta}$, $P_{\alpha\beta}$, and V are undetermined dimensionless constants (except that $P_{\alpha\beta}$ must be such that equations (40) satisfy (35)), $f(\theta, t)$ is an undetermined function of integration, e is the natural logarithm base, and $\delta_{\alpha\beta}$ is the Kronecker delta. Equations (41)–(44) were derived by assuming that the material velocity in a constant stress plastically deforming sector is logarithmically singular, as it must be in the other sector types. Drugan and Chen (1989) have given evidence in support of this assumption; further support derives from the fact that this assumption permits us to construct solutions satisfying all requisite conditions for the present problem.

Now, as noted earlier, solutions to the present plane strain problem of determining the near-tip growing interface crack stress and deformation fields in the ductile material are constructed by assembling angular sectors chosen from the three permissible types summarized previously so that plastic dissipation is non-negative everywhere, the yield condition is not violated in instantaneously elastic sectors, and the following intersector jump conditions (equations (51)–(53)) and problem boundary conditions ((54)–(55)) are satisfied (specializing the forms given in Section 2):

$$[\![\sigma_{\alpha\beta}]\!] = [\![\sigma_{33}]\!] = 0 \tag{51}$$

$$[\![v_\theta]\!] = 0 \tag{52}$$

$$s^\Sigma_{rr} = 0 \text{ and } [\![v_r]\!] = 2V^*[\![\Lambda]\!]\sigma^\Sigma_{r\theta}; \text{ otherwise } [\![v_r]\!] = 0 \tag{53}$$

$$v_1 = v_2 = 0 \quad \text{on } \theta = 0 \tag{54}$$

$$\sigma_{12} = \sigma_{22} = 0 \quad \text{on } \theta = \pi. \tag{55}$$

Since the near-tip boundary conditions do not specify the stress state on the line ahead of the crack, we expect that different near-tip solutions will exist for different ratios of the far-field applied stresses σ_{12}/σ_{22}. Thus, I define a "*near-tip* mixity parameter" M for the plane strain case which is a slight modification of Shih's (1974), as suggested by Shih et al. (1990b) for interface crack problems:

$$M \equiv \frac{2}{\pi}\arctan\left[\lim_{r\to 0}\frac{\sigma_{12}(r,\,\theta=0)}{\sigma_{22}(r,\,\theta=0)}\right]. \tag{56}$$

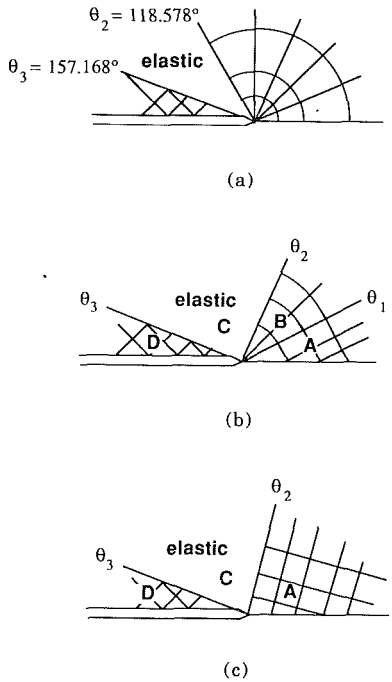


Fig. 3 Near-tip solution configurations in the ductile material for plane strain interfacial crack growth, for each of the two solution families; stress characteristics are shown in plastic sectors. For the "Mode I-type" family, (*a*) illustrates the extreme family member possessing the highest stress triaxiality ($\sigma_{kk}/3 = 5.696k$ on $\theta = 0$; $M = 0.1106$) and a continuous velocity field; (*b*) applies for $0.1106 \geq M \geq -0.1496$ ($0 \leq \theta_1 \leq 72.44$ deg); (c) applies for $-0.1496 \geq M > -0.7362$ (72.44 deg $\leq \theta_1 < 135$ deg). For the "Mode II-type" family, (*a*) with $\theta_2 = 9.706$ deg, $\theta_3 = 179.8$ deg, $M = 0.9449$ is the extreme family member closely resembling the homogeneous Mode II solution and also has continuous velocities; (*b*) applies for $0.9449 \geq M \geq 0.8835$ ($0 \leq \theta_1 \leq 9.862$ deg); (c) applies for $0.8835 \geq M > 0.3325$ (9.862 deg $\leq \theta_1 < 45$ deg).

I shall now show that near-tip growing crack solutions can be found for many values of M in the possible range $-1 \leq M \leq 1$; again, it appears that full-field (numerical) solutions are necessary to determine which value of M corresponds to a given ratio of *far-field* applied stress (as well as load level and specimen geometry).

Before deriving these solutions, we first demonstrate that a subyield sector cannot border $\theta = 0$ by following reasoning identical to that used in the antiplane strain case. The key is boundary condition (54), which requires $A_1 = A_2 = 0$ via equations (49) and (50); this shows from (45)–(47) that the Cartesian stress components in such a potential subyield sector must be constants, but since this sector must border a sector at yield for a nontrivial ($\mathbf{v} \neq \mathbf{0}$) solution, and due to (51), we conclude that these constant stresses must satisfy yield. Therefore, we again find that the sector bordering $\theta = 0$ must be of the centered fan or constant stress plastic type (noting that an elastic sector at yield with $A_1 = A_2 = 0$ may be regarded as a subset of the latter).

A fascinating feature of the present plane strain problem is that *two* near-tip solution families can be found! A key difference between these families is that members of one family have high stress triaxiality levels, involving only hydrostatic *tension* ahead of the growing crack, while members of the second family have much lower triaxialities, with a large range of these latter solutions exhibiting hydrostatic *compression* ahead of the growing crack.

4.2 High Stress Triaxiality Solution Family: Mode I Type

The obvious place to begin a search for near-tip solutions to the present plane strain interfacial crack growth problem is to examine existing solutions for plane strain quasi-static crack growth in *homogeneous* elastic-ideally plastic material. In-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 2 Parameter values for the "Mode I-type" family of plane strain crack growth solutions. Solutions to equations (57)–(58), and values of the parameters appearing in (56), (60), and (65), corresponding to the configuration of Fig. 3(*b*), are given in (*a*); solutions to (67), and values of the parameters appearing in (56), (68), (69), and (65), corresponding to the configuration of Fig. 3(*c*), are given in (*b*).

θ_1 (deg)	θ_2 (deg)	θ_3 (deg)	P_{12}^A	P_{22}^A	P_{11}^A	A_1	A_2	M
0.0	118.57801	157.16784	1.0	5.6960770	5.6960770	−2.6344999	−4.4350645	0.11063714
10.0	118.33631	157.37840	0.93969262	5.6884367	5.0043964	−2.6004145	−4.4023056	0.10422418
20.0	117.57284	158.02831	0.76604444	5.6381857	4.3526104	−2.4988956	−4.3048834	0.08596932
30.0	116.15561	159.17615	0.5	5.5084331	3.7763823	−2.3320335	−4.1452596	0.05762800
40.0	113.79225	160.93484	0.17364818	5.2707412	3.3011257	−2.1028246	−3.9271166	0.02096629
45.0	112.07821	162.10061	0.0	5.1053551	3.1053551	−1.9657674	−3.7973447	0.0
50.0	109.83224	163.50383	−0.17364818	4.9068579	2.9372424	−1.8139916	−3.6542305	−0.02251986
60.0	102.55774	167.25122	−0.5	4.4040210	2.6719702	−1.4641161	−3.3261351	−0.07196894
70.0	84.478480	173.11089	−0.76604444	3.6994770	2.4139017	−1.0212996	−2.9012729	−0.12998673
72.43693	472.436934	175.42170	−0.81788738	3.4158224	2.2650657	−0.86306751	−2.7268311	−0.14961597

(a)

θ_2 (deg)	θ_3 (deg)	P_{12}^A	P_{22}^A	P_{11}^A	A_1	A_2	M
72.436934	175.42170	−0.81788738	3.4158224	2.2650657	−0.86306751	−2.7268311	−0.14961597
80.0	179.79060	−0.93969262	2.3648330	1.6807927	−0.35837076	−2.0324216	−0.24078830
90.0⁻	180.0⁻	−1.0	1.5707963	1.5707963	0.0	−1.5	−0.36090707
100.0	179.99767	−0.93969262	1.0528604	1.7369007	0.19944056	−1.1310836	−0.46388174
110.0	179.98917	−0.76604444	0.64788962	1.9334648	0.28952278	−0.79545731	−0.55307556
120.0	179.99845	−0.5	0.32631750	2.0583683	0.26986240	−0.46741538	−0.63189030
130.0	180.0⁻	−0.17364818	0.08747946	2.0570950	0.12617224	−0.15036622	−0.70291351
135.0⁻	180.0⁻	0.0	0.0	2.0	0.0	0.0	−0.73616961

(b)

deed, scrutiny of the Slepyan (1974)–Rice et al. (1980)–Gao (1980) "modified Prandtl field" Mode I growing crack solution reveals that it solves the present problem in the $M=0$ case, since all near-tip governing equations, intersector continuity conditions and boundary conditions on the ductile side of the interface are the same as in the homogeneous material problem except for the additional requirement here of the first condition of equation (54), which the Mode I solution happens to satisfy. The specifics of this modified Prandtl field solution will be reviewed shortly.

More important than simply providing the solution to the present problem in the $M=0$ case, the modified Prandtl field solution suggests that for a range of M values, its near-tip assembly of sectors, illustrated in Fig. 3(b), may provide a family of growing interfacial crack solutions; this indeed turns out to be the case. As Fig. 3(b) shows, this solution family involves a constant stress plastic sector (A) adjacent to $\theta=0$, followed by a centered fan plastic sector (B) (their border θ_1 depending on M), an elastically deforming sector (C), and finally another constant stress plastic sector (D). Application of the jump and boundary conditions (51)–(55), together with the additional requirement that v_r be continuous across θ_2 (Rice proved (see Drugan and Chen 1989, Appendix A) that the requirements of non-negative plastic work produced by a hypothesized velocity jump and of nonviolation of yield in the elastically deforming sector demand full velocity continuity between a moving centered fan plastic sector and a trailing elastically deforming sector.) to the intrasector field solutions (36)–(50) results in the following pair of coupled equations for θ_2 and θ_3 in terms of a given value of θ_1 (or, equivalently, a given value of shear stress on $\theta=0$):

$$\cos\theta_1\sin\theta_2[2\ln|\sin\theta_2/\sin\theta_3| + \cos2\theta_2 - \cos2\theta_3] - (\cos\theta_1\cos\theta_2 - 1)(\sin2\theta_2 - \sin2\theta_3) - \sin2\theta_2 - 1 = 0 \quad (57)$$

$$\cos\theta_1\sin\theta_2[2(\theta_2-\theta_3) + \sin2\theta_2 - \sin2\theta_3] + (\cos\theta_1\cos\theta_2 - 1)(\cos2\theta_2 - \cos2\theta_3) + \cos2\theta_2 = 0. \quad (58)$$

These equations exhibit physically meaningful solutions for θ_1 in the range $0\le\theta_1\le72.436934$ deg, indicating (see Table 2a) that the configuration of Fig. 3(b) characterizes near-tip solutions for M in the range $0.11063714\ge M\ge-0.14961597$. The near-tip stress and velocity fields for this range of solutions are given by equations (36)–(50), with the parameters appearing in those expressions being given in terms of the three sector boundary angles as follows:

$$V=\cos\theta_1 \quad (59)$$

$$A_1=-3V\sin\theta_2,\quad A_2=3(V\cos\theta_2-1) \quad (60)$$

$$C=S+2\theta_2+(2/3)A_1\ln|\sin\theta_2/\sin\theta_3|+(2/3)A_2(\theta_2-\theta_3) \quad (61)$$

$$C_{11}=6S-4A_1\ln|\sin\theta_3|-A_1\cos2\theta_3-A_2(2\theta_3+\sin2\theta_3) \quad (62)$$

$$C_{22}=A_1\cos2\theta_3-A_2(2\theta_3-\sin2\theta_3) \quad (63)$$

$$C_{12}=-A_1(2\theta_3+\sin2\theta_3)+A_2\cos2\theta_3 \quad (64)$$

$$P_{12}^A=\cos2\theta_1,\ P_{22}^A=\sin2\theta_1+C-2\theta_1,\ P_{11}^A=-\sin2\theta_1+C-2\theta_1,\ B_\alpha^A=0 \quad (65)$$

$$P_{11}^D=2S,\ P_{12}^D=P_{22}^D=0,\ B_\alpha^D=A_\alpha. \quad (66)$$

The parameter S is simply the sign of P_{11}^D: $S=+1$ here; it is introduced now for later convenience. Superscripts A and D denote parameters valid in Sectors A and D, respectively, of Fig. 3(b). Solutions to equations (57)–(58) for θ_1 in the ad-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

missible range $0 \le \theta_1 \le 72.436934$ deg are given in Table 2*a* together with the values of M, the Cartesian components of stress in the leading constant stress plastic sector normalized by k, and the leading-order velocity amplitude coefficients for Sectors C and D, via (56), (65), and (60), respectively. Note first that the entries in the row having $\theta_1 = 45$ deg correspond to the homogeneous Mode I modified Prandtl field solution mentioned previously. One then observes that all solutions in Table 2*a* for θ_1 in the range $0 \le \theta_1 < 45$ deg (i.e., M in the range $0.11063714 \ge M > 0$) involve higher stress triaxiality on the line ahead of the crack than the Mode I homogeneous solution, with the highest triaxiality level being about 39 percent higher! This occurs for the extreme family member having $\theta_1 = 0$ deg, which is illustrated in Fig. 3(*a*). Thus, one clear effect of the presence of a stiff brittle material on one side of the crack plane is an elevated crack-line stress triaxiality for a substantial range of shear/normal stress ratios, as compared to homogeneous Mode I or Mode II (to be discussed) solutions. It is also worth mentioning that just as the homogeneous Mode I modified Prandtl field does, all the solutions just presented, except the limiting case shown in Fig. 3(*a*), involve zero velocities (at least to first order in r) in the leading constant stress plastic sector and hence a radial velocity discontinuity across θ_1, which is permitted by the jump condition (53) and because it produces positive plastic work.

It was noted previously that the configuration of Fig. 3(*b*) characterizes physically meaningful solutions for θ_1 in the range $0 \le \theta_1 \le 72.436934$ deg, and the results in Table 2*a* illustrate why this is: As $\theta_1 \to 72.436934$ deg, $\theta_2 \to \theta_1$. Thus, the centered fan vanishes as this angle is approached suggesting, via behavior remarkably similar to the antiplane family discussed in Section 3, that further solutions can be found having a simpler "fanless" configuration. This is again the case. Figure 3(*c*) illustrates this configuration, which involves a leading constant stress plastic sector (A), followed by an elastically deforming sector (C) which borders a trailing constant stress plastic sector (D). Application of conditions (51)–(55) to (40)–(50) for this configuration leads to the following results (θ_2 must be a line of radial velocity discontinuity for solutions of this configuration to be found; equations (53) and (35) thus require $|\sigma_{r\theta}(\theta_2)| = k$, i.e., θ_2 must be a stress characteristic):

$$(\cos\theta_2/\sin\theta_2)[2\cos 2\theta_2 \ln|\sin\theta_2/\sin\theta_3| + 2(\theta_2 - \theta_3)(S + \sin 2\theta_2) + S(\sin 2\theta_2 - \sin 2\theta_3) + 1 - \cos 2(\theta_2 - \theta_3)] - S(\cos 2\theta_2 - \cos 2\theta_3) + \sin 2(\theta_2 - \theta_3) = 0 \tag{67}$$

$$A_2 = \frac{3\cos 2\theta_2}{(\cos\theta_2/\sin\theta_2)[2(\theta_2 - \theta_3) + \sin 2\theta_2 - \sin 2\theta_3] - \cos 2\theta_2 + \cos 2\theta_3} \tag{68}$$

$$A_1 = A_2(\cos\theta_2/\sin\theta_2) \tag{69}$$

with the other parameters in (40)–(50) being given by (61)–(66), setting $\theta_1 = \theta_2$ in (65). Equation (67) has physically meaningful solutions for θ_2 in the range $72.436934 \text{ deg} \le \theta_2 < 135$ deg; these are displayed in Table 2*b*, which shows that the corresponding range for M is $-0.14961597 \ge M > -0.73616961$. For $72.436934 \text{ deg} \le \theta_2 \le 90$ deg, these solutions have a *tensile* σ_{11} in the trailing constant stress plastic sector, so $S = +1$, whereas for $90 \text{ deg} < \theta_2 < 135$ deg, σ_{11} is *compressive* in that sector, meaning that $S = -1$. It is interesting to note that (67) does have solutions for $\theta_2 < 72.436934$ deg, but these solutions can be shown to violate the yield condition in the elastically deforming sector of Fig. 3(*c*) (except for those to be presented in Section 4.3) and are thus physically meaningless.

Table 2*b* shows that as $\theta_2 \to 135$ deg, both A_1 and $A_2 \to 0$, meaning that near-tip velocities vanish everywhere in this limit and thus signaling the termination of this tensile-type interfacial crack growth family. As Tables 2 show, the family spans an M range of $-0.73616961 < M \le 0.11063714$, and a crack-line stress triaxiality range of $k < \sigma_{kk}/3 \le 5.6960770k$. Also observe that the crack-line shear stress sweeps continuously through its entire range of admissible values (delineated by equation (35)), from $k \ge \sigma_{12}(\theta = 0) \ge -k$, and again through the negative half of this range, as M decreases from its highest to lowest value in this family.

4.3 Low Stress Triaxiality Solution Family: Mode II-Type. We have already observed that the Mode I modified Prandtl field growing crack solution for *homogeneous* material is also a solution (in the ductile material) to the present interfacial crack growth problem for the case $M = 0$. In fact, the solution family presented in Section 4.2 was determined by beginning with this solution and seeking other solutions as θ_1 of Fig. 3(*b*) was permitted to vary from 45 deg. It thus seemed worthwhile also to examine the *homogeneous* material Mode II growing crack solution, which is due to Slepyan (1974). This solution has the same configuration as that illustrated in Fig. 3(*a*), but with sector boundary angles of $\theta_2 = 9.5661$ deg, $\theta_3 = 179.820$ deg. However, this is *not* a solution to the present interface crack growth problem because the second of boundary conditions (54) is violated. Nevertheless, perhaps because the relative magnitude of the violation of this boundary condition is small, it seemed advisable to investigate whether a solution to the present problem could be found that closely resembles the homogeneous material Mode II Slepyan solution. This amounts to seeking another solution to (the highly nonlinear) equations (57)–(58) for $\theta_1 = 0$. Remarkably, such a solution does exist; it is listed in the first row of Table 3*a*, and one observes that it indeed differs very little from the homogeneous material Slepyan solution quoted above. Having just discovered the solution family described in Section 4.2, the existence of this Mode II-type solution, which has the same geometrical configuration (but vastly different sector angular extents!) as the extreme Mode I-type family member illustrated in Fig. 3(*a*), raised the interesting possibility of another solution family having the same near-tip sector configurations as this Mode I-type family. Such a second solution family does indeed exist: It represents a second set of solutions to equations (57)–(58) for the first part of the family, and to equation (67) for the second part of the family. Again, it is the nonlinearity of the governing equations that permits these two solution families.

Thus, the first part of this Mode II-type solution family has a near-tip configuration identical to that of Fig. 3(*b*), and hence the governing equations (57)–(58) apply. The near-tip stress and velocity fields are given by (36)–(50), with the parameters appearing in those expressions again being given in terms of the three sector boundary angles by (59)–(66) with $S = +1$. This family portion has solutions for θ_1 in the narrow range $0 \le \theta_1 \le 9.8623655$ deg, indicating that the configuration of Fig. 3(*b*) characterizes near-tip solutions for M in the range $0.94493883 \ge M \ge 0.88348201$; these solutions are presented in Table 3*a*. They show that as $\theta_1 \to 9.8623655$ deg, $\theta_2 \to \theta_1$, signaling, just as in the Mode I-type family of Section 4.2, a changeover to the simpler near-tip configuration of Fig. 3(*c*). Thus, for this second part of the present family, the governing equation (67) applies, and the near-tip stress and velocity fields are given by (40)–(50) with the parameters in those expressions being given in terms of the two sector boundary angles by (68), (69), and (61)–(66), setting $\theta_1 = \theta_2$ in (65). Here again, $S = +1$. Equation (67) admits this second set of physically meaningful solutions for θ_2 in the range $9.8623655 \text{ deg} \le \theta_2 < 45$ deg, indicating that Fig. 3(*c*) characterizes near-tip solutions for M in the range $0.88348201 \ge M > 0.33249070$; these solutions are presented in Table 3*b*. Again, similar to the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 3 Parameter values for the "Mode II-type" family of plane strain crack growth solutions. Since this family is another set of solutions to the same equations as the "Mode I-type" family, (*a*) and (*b*) here present the same information as they do in Table 2, but for this second solution family.

θ_1 (deg)	θ_2 (deg)	θ_3 (deg)	P_{12}^A	P_{22}^A	P_{11}^A	A_1	A_2	M
0.00000000	9.7057530	179.81688	1.0	0.08670619	0.08670619	−0.50576506	−0.04294036	0.94493883
5.0	9.7455982	179.81598	0.98480775	0.11009657	−0.23719978	−0.50588895	−0.05454396	0.92912340
9.8623655	9.8623655	179.81336	0.94132495	0.17423689	−0.50076631	−0.50625240	−0.08801257	0.88348201

(a)

θ_2 (deg)	θ_3 (deg)	P_{12}^A	P_{22}^A	P_{11}^A	A_1	A_2	M
9.8623655	179.81336	0.94132495	0.17423689	−0.50076631	−0.50625240	−0.08801257	0.88348201
15.0	179.83743	0.86602540	0.25641648	−0.74358352	−0.49844690	−0.13355844	0.81674231
20.0	179.89182	0.76604444	0.32110427	−0.96447094	−0.47388461	−0.17247989	0.74731045
25.0	179.94814	0.64278761	0.36245635	−1.1696325	−0.43098790	−0.20097296	0.67313480
30.0	179.98586	0.5	0.37043832	−1.3616125	−0.36762002	−0.21224552	0.59406797
35.0	179.99898	0.34202014	0.33109897	−1.5482863	−0.28027021	−0.19624731	0.51032808
40.0	180.0⁻	0.17364818	0.22191517	−1.7477003	−0.16235262	−0.13623003	0.42270122
45.0⁻	180.0⁻	0.0	0.0	−2.0	0.0	0.0	0.33249070

(b)

Mode I-type family, equation (67) does have solutions for $\theta_2 < 9.8623655$ deg, but these solutions violate the yield condition in the elastically deforming sector of Fig. 3(*c*) and hence are physically inadmissible.

Observe from Table 3*b* that as $\theta_2 \to 45$ deg, both A_1 and $A_2 \to 0$, meaning that near-tip velocities vanish everywhere in this limit and thus signaling the termination of this shear-type interfacial crack growth family. As Tables 3 show, the family spans an M range of $0.33249070 < M \leq 0.94493883$, and a crack-line stress triaxiality range of $-k < \sigma_{kk}/3 \leq 0.08670619k$. Also, observe that the crack-line shear stress sweeps continuously through its entire range of admissible positive values, $k \geq \sigma_{12}(\theta = 0) > 0$, as M decreases from its highest to lowest value in this family.

4.4 Discussion. Observing that $B_2^D = A_2 < 0$ in all of the solutions presented here (see Tables 2 and 3), equation (42) shows that all solutions correspond to an *opening* near-tip crack profile during growth. It appears that the results presented here represent all possible such solutions. In support of this statement I observe: that neither of the two solution families presented can be extended further without a region of negative plastic work production; that for these two families together, the crack-line shear stress sweeps continuously through its entire admissible range of values two times; and that for various reasons cited throughout the analysis, other near-tip sector combinations seem highly unlikely. Thus, an interesting conclusion of this study is that for plane strain ductile/brittle interfacial crack growth under the assumptions of Section 2, solutions involving near-tip opening of the growing crack are possible only for near-tip mixities lying within the ranges $-0.73616961 < M \leq 0.11063714$ and $0.33249070 < M \leq 0.94493883$.

Although the experimental evidence seems as yet to be too sketchy to suggest or confirm a crack growth criterion for ductile/brittle interface cracks, observe that the near-tip solutions derived herein can be employed to develop a quantitative crack growth criterion based, e.g., on geometrical self-similarity of the propagating crack tip in the manner of Rice et al. (1980).

Finally, it is interesting to notice that the problem analyzed here does not admit families of near-tip solutions of the type derived by Drugan and Chen (1989), which have curved intersector boundaries and fully continuous near-tip velocity fields. Solution families of that type apparently require nonzero material velocity on $\theta = 0$, which is ruled out here by the assumption that the brittle material is *rigid* (and bonding is perfect). That is, the assumption that the brittle phase is rigid appears to give rise to solutions with velocity discontinuities (except for the two extreme solutions in which a centered fan borders $\theta = 0$.) Another implication, which seems to be in accord with physical reasoning, is that if one material phase is idealized as being truly rigid while the other is incompressible elastic-ideally plastic, near-tip growing crack fields seem unable to display the effects of yielding extent, specimen geometry or load level except through changes of local mixity M. Alternatively, if the brittle material is treated as elastic (even with very high stiffness) rather than rigid, it appears that solution families of the Drugan and Chen (1989) type may exist, and when the elastic phase has very high stiffness, extreme members of these families (corresponding to "small-scale yielding" situations) should resemble very closely the solutions derived in the present work, except that they will have fully continuous velocity fields at all mixities for which solutions exist. We are currently investigating these hypotheses. The present study should serve as a useful guide to such more difficult analyses.

Addendum

At the same meeting at which the present paper was delivered (see footnote to title), Ponte Castaneda and Mataga (1990) presented results on near-tip fields for stable ductile/brittle interfacial crack growth, in which the ductile material was characterized by J_2 flow theory with either linear hardening or perfect plasticity. In the perfect plasticity case, they treated the brittle material as rigid and sought solutions in

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the ductile material under the restrictive assumption of fully continuous velocity fields (as opposed to the present work, where I allow propagating surfaces of tangential velocity discontinuity, as permitted by the Drugan and Rice (1984) moving discontinuity analysis); the discrete solutions they found coincide with extreme members of the solution families derived here, which do have fully continuous velocity fields and are characterized by having a centered fan plastic sector bordering the crack plane ahead of the crack, as detailed in the foregoing.

Acknowledgments

This work was supported by the Solid Mechanics Program of the U.S. National Science Foundation under Grant MSM-8552486, together with a matching grant from Rockwell International, and by the Mechanics Division of the U.S. Office of Naval Research under Grant N0014-89-J-1206. I thank Dr. Xing-Yu Chen for performing some preliminary numerical calculations.

References

Chitaley, A. D., and McClintock, F. A., 1971, "Elastic-Plastic Mechanics of Steady Crack Growth Under Anti-Plane Shear," *Journal of the Mechanics and Physics of Solids*, Vol. 19, pp. 147–163.

Clements, D. L., 1971, "A Crack Between Dissimilar Anisotropic Media," *International Journal of Engineering Science*, Vol. 9, pp. 257–265.

Drugan, W. J., 1985, "On the Asymptotic Continuum Analysis of Quasistatic Elastic-Plastic Crack Growth and Related Problems," ASME Journal of Applied Mechanics, Vol. 52, pp. 601–605.

Drugan, W. J., 1986, "A More Direct and General Analysis of Moving Strong Discontinuity Surfaces in Quasi-Statically Deforming Elastic-Plastic Solids," ASME Journal of Applied Mechanics, Vol. 53, pp. 224–226.

Drugan, W. J., and Chen, X.-Y., 1989, "Plane Strain Elastic-Ideally Plastic Crack Fields for Mode I Quasistatic Growth at Large-Scale Yielding I. A New Family of Analytical Solutions," *Journal of the Mechanics and Physics of Solids*, Vol. 37, pp. 1–26.

Drugan, W. J., and Rice, J. R., 1984, "Restrictions on Quasi-Statically Moving Surfaces of Strong Discontinuity in Elastic-Plastic Solids," *Mechanics of Material Behavior: The D. C. Drucker Anniversary Volume*, G. J. Dvorak and R. T. Shield, eds., Elsevier, Amsterdam, pp. 59–73.

England, A. H., 1965, "A Crack Between Dissimilar Media," ASME Journal of Applied Mechanics, Vol. 32, pp. 400–402.

Erdogan, F., 1963, "Stress Distribution in a Nonhomogeneous Elastic Plane with Cracks," ASME Journal of Applied Mechanics, Vol. 30, pp. 232–236.

Erdogan, F., 1965, "Stress Distribution in Bonded Dissimilar Materials with Cracks," ASME Journal of Applied Mechanics, Vol. 32, pp. 403–410.

Gao, Y. C., 1980, "Elastic-Plastic Field at the Tip of a Crack Growing Steadily in Perfectly Plastic Medium," (in Chinese) *Acta Mechanica Sinica*, No. 1, pp. 48–56.

Hutchinson, J. W., Mear, M. E., and Rice, J. R., 1987, "Crack Paralleling an Interface Between Dissimilar Materials," ASME Journal of Applied Mechanics, Vol. 54, pp. 828–832.

Needleman, A., 1990a, "An Analysis of Decohesion Along an Imperfect Interface," *International Journal of Fracture*, Vol. 42, pp. 21–40.

Needleman, A., 1990b, "An Analysis of Tensile Decohesion Along an Interface," *Journal of the Mechanics and Physics of Solids*, Vol. 38, pp. 289–324.

Park, J. H., and Earmme, Y. Y., 1986, "Application of Conservation Integrals to Interfacial Crack Problems," *Mechanics of Materials*, Vol. 5, pp. 261–276.

Parks, D. M., and Zywicz, E., 1989, "Elastic/Perfectly Plastic Small Scale Yielding at Bi-Material Interfaces," *Advances in Fracture Research, Proceedings of ICF7*, K. Salama et al., eds., Vol. 4, Pergamon Press, Oxford, pp. 3081–3088.

Ponte Castaneda, P., and Mataga, P. A., 1990, "Stable Crack Growth Along a Brittle/Ductile Interface, Part I–Near-Tip Fields," submitted for publication.

Rice, J. R., 1982, "Elastic-Plastic Crack Growth," *Mechanics of Solids: The R. Hill 60th Anniversary Volume*, H. G. Hopkins and M. J. Sewell, eds., Pergamon Press, Oxford, pp. 539–562.

Rice, J. R., 1988, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," ASME Journal of Applied Mechanics, Vol. 55, pp. 98–103.

Rice, J. R., Drugan, W. J., and Sham, T. L., 1980, "Elastic-Plastic Analysis of Growing Cracks," *Fracture Mechanics: Twelfth Conference*, ASTM STP 700, American Society for Testing and Materials, Philadelphia, Pa., pp. 189–221.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," ASME Journal of Applied Mechanics, Vol. 32, pp. 418–423.

Shih, C. F., 1974, "Small-Scale Yielding Analysis of Mixed-Mode Plane-Strain Crack Problems," *Fracture Analysis*, ASTM STP 560, ASTM, Philadelphia, Pa., pp. 187–210.

Shih, C. F., and Asaro, R. J., 1988, "Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces: Part I—Structure of Small Scale Yielding Fields," ASME Journal of Applied Mechanics, Vol. 55, pp. 299–316.

Shih, C. F., and Asaro, R. J., 1989a, "Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces: Part II—Structure of Small Scale Yielding Fields," ASME Journal of Applied Mechanics, Vol. 56, pp. 763–779.

Shih, C. F., and Asaro, R. J., 1989b, "Elastic-Plastic Analysis of a Collinear Array of Cracks on a Bimaterial Interface," *Materials Science and Engineering*, Vol. A107, pp. 145–157.

Shih, C. F., Asaro, R. J., and O'Dowd, N. P., 1990a, "Cracks on Bimaterial Interfaces: Plasticity Aspects," *Proceedings of the Acta/Scripta Metallurgica Conference on Metal/Ceramic Interfaces*, Series 4, pp. 313–325.

Shih, C. F., Asaro, R. J., and O'Dowd, N. P., 1990b, "Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces: Part III—Large Scale Yielding," ASME Journal of Applied Mechanics, in press.

Sih, G. C., and Rice, J. R., 1964, "The Bending of Plates of Dissimilar Materials with Cracks," ASME Journal of Applied Mechanics, Vol. 31, pp. 477–482.

Slepyan, L. I., 1974, "Growing Crack During Plane Deformation of an Elastic-Plastic Body," *Mekhanika Tverdogo Tela*, Vol. 9, pp. 57–67.

Varias, A. G., O'Dowd, N. P., Asaro, R. J., and Shih, C. F., 1990, "Crack Growth on Bimaterial Interfaces," *Proceedings of the Acta/Scripta Metallurgica Conference on Metal/Ceramic Interfaces*, Series 4, pp. 375–382.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bulletin of the Seismological Society of America*, Vol. 49, pp. 199–204.

Willis, J. R., 1971, "Fracture Mechanics of Interface Cracks," *Journal of the Mechanics and Physics of Solids*, Vol. 19, pp. 353–368.

Zywicz, E., and Parks, D. M., 1990, "Elastic-Plastic Analysis of Frictionless Contact at Interfacial Crack Tips," *International Journal of Fracture*, Vol. 42, pp. 129–144.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Li Yingzhi[1]

D. A. Hills

Department of Engineering Science,
University of Oxford,
Oxford, OX1 3PJ, U.K.

The Hertzian Cone Crack

Stress intensity factors are determined for cracks in the form of a frustum of a cone, and loaded by a Hertzian contact pressure distribution using a global-local finite element solution. These will be helpful in interpreting blunt indentation tests quantitatively, and it is shown that for most crack geometries occurring in practice, the mode II component is about a tenth of the mode I contribution.

Introduction

It has long been observed that if a steel ball is steadily pressed into a large, flat glass block, a cone-shaped crack forms suddenly at a characteristic load (Frank and Lawn (1967)). However, although the phenomenon was first observed almost a century ago, a quantitative treatment of the effect has remained elusive. It is clear that the main propelling force is the contact stress field, which is approximately Hertzian, if the contacting surfaces are assumed smooth. This was first analyzed by Hüber (1904), and a solution is available in closed form. However, it is not possible to explain the experimental result that the fracture load is almost proportional to the radius of the sphere (Auerbach's law) by consideration of this stress state alone. There are two complicating effects: First, the difference between the elastic constants of steel and glass causes significant shear tractions to arise, and this facilitated an explanation of the observation that cracks must initiate outside the contact disk, since this is where the direct stress lying in the surface reaches a maximum (Johnson et al., 1973). Encouraged by this success, a fuller treatment of the resulting stress field was carried out (Hills and Sackfield, 1987).

The second complication is that propagation of the crack is not stable, and hence there is a spontaneous increase in crack size. This may be treated using fracture mechanics, but expressions for the stress intensity factors used by Frank and Lawn (1967) were for plane cracks, and a quantitative explanation requires a knowledge of the stress intensity for axisymmetric conical cracks. There has been only a limited amount of success in solving such a configuration. For instance, Yashi and Parnas (1986) considered a crack in the form of a frustum of a cone subject to uniform tractions, and none has dealt with their response in a Hertzian field. However, results have been published for a short cylindrical crack propelled by a Hertzian contact by Finnie et al. (1981).

The intention of the present paper is to fill this gap, by providing the missing stress intensity factors, and hence enabling a quantitative assessment of the material's fracture toughness to be made. Glass is the ideal material in which to carry out these tests, since the cone angle may be readily observed. However, glass is of limited engineering interest, and indentation testing may be applied to ceramics. Results are presented for $\nu = 0.23$ (glass) and $\nu = 0.3$.

The technique to be used is the so-called "local-global analysis," whereby a special finite element is devised giving the correct asymptotic behavior at the crack tip. An important pioneering paper addressing a similar problem, brought to the authors' attention by a referee, is by Zak (1972).

The method of attack is as follows:

(1) Introduce a special coordinate transformation with parameters ρ, θ, ϕ to enable the crack boundary conditions to be specified simply.

(2) Establish the governing equations in curvilinear coordinates. Use the eigenexpansion method to find a set of eigenfunctions which satisfies the crack boundary condition.

(3) Establish a high-order axisymmetric special finite element, based on the stress-strain fields at crack tip to determine the unknown coefficients in the expansion and yield the stress intensity factors directly.

(4) Solve the combined contact/crack problem by finite elements incorporating the special crack-tip element to accelerate convergence.

Governing Equations in Curvilinear Coordinates

First, introduce a special set of coordinates ρ, θ, ϕ, Fig. 1, to permit simplified crack boundary conditions to be specified along $\theta = \pm\pi$. The relation between the Cartesian coordinate set and the new one is

$$x = [a + l\sin\theta_0 + \rho\sin(\theta + \theta_0)]\cos\phi$$
$$y = [a + l\sin\theta_0 + \rho\sin(\theta + \theta_0)]\sin\phi$$
$$z = l\cos\theta_0 + \rho\cos(\theta + \theta_0). \quad (1)$$

From equation (1) the metric tensor and Christoffel symbols can be calculated. From these, the equations of equilibrium together with Hooke's law and strain displacement were established (Li and Hills, 1988). For brevity we will restrict ourselves here to the case of axisymmetric loading, so the displacement in the circumferential direction, u_ϕ vanishes, and hence quantities depending on its derivatives, i.e., $\epsilon_{\rho\phi}$, $\epsilon_{\theta\phi}$ also disappear, together with corresponding stress components. The

[1]Permanent address: Institute of Mechanics, Chinese Academy of Sciences, People's Republic of China.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 8, 1988; final revision, Jan. 11, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

resulting equations of equilibrium are, expressed in physical (rather than tensor) quantities,

$$\frac{\partial\sigma_{\rho\rho}}{\partial\rho}+\frac{\partial\sigma_{\rho\theta}}{\rho\partial\theta}+\frac{\sigma_{\rho\rho}-\sigma_{\theta\theta}}{\rho}+$$
$$+\frac{1}{\Phi}[(\sigma_{\rho\rho}-\sigma_{\phi\phi})\sin(\theta+\theta_0)+\sigma_{\rho\theta}\cos(\theta+\theta_0)]=0,$$
$$\frac{\partial\sigma_{\rho\theta}}{\partial\rho}+\frac{\partial\sigma_{\theta\theta}}{\rho\partial\theta}+\frac{2\sigma_{\theta\rho}}{\rho}+$$
$$+\frac{1}{\Phi}[\sigma_{\rho\theta}\sin(\theta+\theta_0)+(\sigma_{\theta\theta}-\sigma_{\phi\phi})\cos(\theta+\theta_0)]=0 \qquad (2)$$

where $\Phi=a+l\sin\theta_0+\rho\sin(\theta+\theta_0)$. It represents Lamé's coefficient in the ϕ-direction. The geometric or strain-displacement relations are, again in terms of physical quantities,

$$\epsilon_{\rho\rho}=\frac{\partial u_\rho}{\partial\rho}$$
$$\epsilon_{\theta\theta}=\frac{\partial u_\theta}{\rho\partial\theta}+\frac{u_\rho}{\rho}$$
$$\epsilon_{\phi\phi}=\frac{1}{\Phi}[u_\rho\sin(\theta+\theta_0)+u_\theta\cos(\theta+\theta_0)]$$
$$\epsilon_{\rho\theta}=\frac{\partial u_\theta}{\partial\rho}+\frac{\partial u_\rho}{\rho\partial\theta}-\frac{u_\theta}{\rho} \qquad (3)$$

and we have $\epsilon_{\rho\phi}=0$, $\epsilon_{\theta\phi}=0$.

Hooke's law is

$$\sigma_{\rho\rho}=\lambda\Theta+2\mu\frac{\partial u_\rho}{\partial\rho}$$
$$\sigma_{\theta\theta}=\lambda\Theta+2\mu\left(\frac{\partial u_\theta}{\rho\partial\theta}+\frac{u_\rho}{\rho}\right)$$
$$\sigma_{\phi\phi}=\lambda\Theta+2\mu\frac{1}{\Phi}[u_\rho\sin(\theta+\theta_0)+u_\theta\cos(\theta+\theta_0)] \qquad (4)$$
$$\sigma_{\rho\theta}=\mu\left(\frac{\partial u_\theta}{\partial\rho}+\frac{\partial u_\rho}{\rho\partial\theta}-\frac{u_\theta}{\rho}\right)$$

where the dilation Θ is given by

$$\Theta=\frac{\partial u_\rho}{\partial\rho}+\frac{\partial u_\theta}{\rho\partial\theta}+\frac{u_\rho}{\rho}+\frac{1}{\Phi}[u_\rho\sin(\theta+\theta_0)+u_\theta\cos(\theta+\theta_0)]$$

and $\sigma_{\rho\phi}=0$, $\sigma_{\theta\phi}=0$.

Substituting equations (3) and (4) into equation (2), we arrive at the following governing equations for the problem, expressed in the form of an unknown displacement field.

$$2(1-\nu)\left(\frac{\partial^2 u_\rho}{\partial\rho^2}+\frac{\partial u_\rho}{\rho\partial\rho}-\frac{u_\rho}{\rho^2}\right)+(1-2\nu)\frac{1}{\rho^2}\frac{\partial^2 u_\rho}{\partial\theta^2}$$
$$+\frac{1}{\rho}\frac{\partial^2 u_\theta}{\partial\rho\partial\theta}-(3-4\nu)\frac{1}{\rho^2}\frac{\partial u_\theta}{\partial\theta}$$
$$+\frac{1}{\Phi}\left[2(1-\nu)\frac{\partial u_\rho}{\partial\rho}\sin(\theta+\theta_0)+\frac{\partial u_\theta}{\partial\rho}\cos(\theta+\theta_0)\right.$$
$$\left.+(1-2\nu)\left(\frac{\partial u_\rho}{\rho\partial\theta}-\frac{u_\theta}{\rho}\right)\cos(\theta+\theta_0)\right]$$
$$-\frac{2}{\Phi^2}(1-\nu)[u_\rho\sin(\theta+\theta_0)+u_\theta\cos(\theta+\theta_0)]\sin(\theta+\theta_0)=0 \qquad (5a)$$

$$\frac{\partial^2 u_\rho}{\rho\partial\rho\partial\theta}+(3-4\nu)\frac{1}{\rho^2}\frac{\partial u_\rho}{\partial\theta}+(1-2\nu)\left(\frac{\partial^2 u_\theta}{\partial\rho^2}+\frac{\partial u_\theta}{\rho\partial\rho}-\frac{u_\theta}{\rho^2}\right)$$
$$+2(1-\nu)\frac{1}{\rho^2}\frac{\partial^2 u_\theta}{\partial\theta^2}+\frac{1}{\Phi}\left[\frac{\partial u_\rho}{\rho\partial\theta}\sin(\theta+\theta_0)\right.$$
$$+2(1-\nu)\frac{u_\rho}{\rho}\cos(\theta+\theta_0)+2(1-\nu)\frac{\partial u_\theta}{\rho\partial\theta}\cos(\theta+\theta_0)$$
$$\left.-\frac{u_\theta}{\rho}\sin(\theta+\theta_0)+(1-2\nu)\frac{\partial u_\theta}{\partial\rho}\sin(\theta+\theta_0)\right]$$
$$-\frac{2}{\Phi^2}(1-\nu)[u_\rho\sin(\theta+\theta_0)+u_\theta\cos(\theta+\theta_0)]\cos(\theta+\theta_0)=0 \qquad (5b)$$

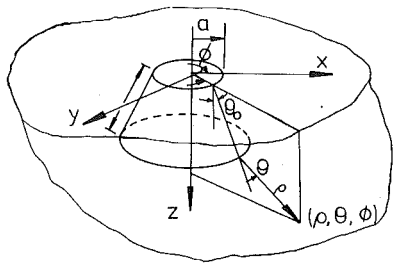


Fig. 1 Coordinate transformation

We shall assume that the crack lies in a nominally tensile region so that it is open throughout its length. The requirement that the crack faces remain free of traction is therefore

$$\sigma_{\theta\theta}=0,\ \sigma_{\rho\theta}=0 \quad \text{when } \theta=\pm\pi. \qquad (6)$$

The other shear traction $\sigma_{\theta\phi}$ clearly vanishes from symmetry.

Eigenfunction Expansion

It is our intention to recast the governing equations in terms of dimensionless displacements, and then to develop a double eigenexpansion series, valid in the neighborhood of the crack tip. This technique was developed by Williams (1957) and has been used to solve a range of crack problems including, recently, some in complex heterogeneous stress fields (Liu and Li, 1987; Li, 1988). First, the Lamé's coefficient in the ϕ-direction is used to introduce a normalization variable. For small ρ, i.e., near the crack tip, we may write

$$\frac{1}{\Phi}=\frac{1}{a+l\sin\theta_0+\rho\sin(\theta+\theta_0)}=\frac{1}{R}[1-r\sin(\theta+\theta_0)+.....] \qquad (7)$$

where R is the radius of the base of the cone frustum and

$$R=a+l\sin\theta_0$$
$$r=\rho/R.$$

Dimensionless displacements are then defined by

$$v_r=u_\rho/R,$$
$$v_\theta=u_\theta/R. \qquad (8)$$

Using the expression for $1/\Phi$ contracted to the first two terms of the expansion the governing equations (5) may now be rewritten in terms of v_r, v_θ as

$$2(1-\nu)\left(\frac{\partial^2 v_r}{\partial r^2}+\frac{\partial v_r}{r\partial r}-\frac{v_r}{r^2}\right)+(1-2\nu)\frac{1}{r^2}\frac{\partial^2 v_r}{\partial\theta^2}+\frac{1}{r}\frac{\partial^2 v_\theta}{\partial r\partial\theta}$$
$$-(3-4\nu)\frac{1}{r^2}\frac{\partial v_\theta}{\partial\theta}+[1-r\sin(\theta+\theta_0)+\ldots][2(1-\nu)\frac{\partial v_r}{\partial r}\sin(\theta+\theta_0)$$
$$+\frac{\partial v_\theta}{\partial r}\cos(\theta+\theta_0)+(1-2\nu)\left(\frac{\partial v_r}{r\partial\theta}-\frac{v_\theta}{r}\right)\cos(\theta+\theta_0)]$$
$$-2(1-\nu)[1-2r\sin(\theta+\theta_0)+\ldots][v_r\sin(\theta+\theta_0)$$
$$+v_\theta\cos(\theta+\theta_0)]\sin(\theta+\theta_0)=0 \qquad (9a)$$

$$\frac{\partial^2 v_r}{r\partial r\partial\theta}+(3-4\nu)\frac{1}{r^2}\frac{\partial v_r}{\partial\theta}+(1-2\nu)\left(\frac{\partial^2 v_\theta}{\partial r^2}+\frac{\partial v_\theta}{r\partial r}-\frac{v_\theta}{r^2}\right)$$
$$+2(1-\nu)\frac{1}{r^2}\frac{\partial^2 v_\theta}{\partial\theta^2}+[1-r\sin(\theta+\theta_0)+\ldots]\left[\frac{\partial v_r}{r\partial\theta}\sin(\theta+\theta_0)\right.$$
$$\left.+2(1-\nu)\left(\frac{\partial v_\theta}{r\partial\theta}+\frac{v_r}{r}\right)\cos(\theta+\theta_0)\right.$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-\frac{v_\theta}{r}\sin(\theta+\theta_0)+(1-2\nu)\frac{\partial v_\theta}{\partial r}\sin(\theta+\theta_0)\Bigg]$$
$$-2(1-\nu)[1-2r\sin(\theta+\theta_0)+\ldots][v_r\sin(\theta+\theta_0)$$
$$+v_\theta\cos(\theta+\theta_0)]\cos(\theta+\theta_0)=0. \quad (9b)$$

Similarly, the boundary conditions may be expressed in terms of dimensionless displacements, using equations (6), (4), and (8) giving

$$\sigma_{\theta\theta}=\frac{E(1-\nu)}{(1+\nu)(1-2\nu)}\left(\frac{\partial v_\theta}{r\partial\theta}+\frac{v_r}{r}+\frac{\nu}{1-\nu}\frac{\partial v_r}{\partial r}\right)+\frac{E\nu}{(1+\nu)(1-2\nu)}$$
$$\times\ [1-r\sin(\theta+\theta_0)+\ldots][v_r\sin(\theta+\theta_0)+v_\theta\cos(\theta+\theta_0)]=0 \quad (10a)$$

$$\sigma_{\rho\theta}=\frac{E}{2(1+\nu)}\left(\frac{\partial v_\theta}{\partial r}+\frac{\partial v_r}{r\partial\theta}-\frac{v_\theta}{r}\right)=0 \quad \text{when } \theta=\pm\pi. \quad (10b)$$

Following Williams (1957), we now expand the dimensionless displacements v_r, v_θ as a double eigenexpansion series,

$$v_r=\sum_\lambda\sum_n r^{\lambda+n}a_n(\theta,\lambda)=\sum_\lambda r^\lambda[a_0(\theta,\lambda)+ra_1(\theta,\lambda)+\ldots]$$
$$v_\theta=\sum_\lambda\sum_n r^{\lambda+n}b_n(\theta,\lambda)=\sum_\lambda r^\lambda[b_0(\theta,\lambda)+rb_1(\theta,\lambda)+\ldots]. \quad (11)$$

Substituting the leading terms for v_r, v_θ into governing equations (9) yields the following pair of equations

$$2(1-\nu)(\lambda^2-1)a_0+(1-2\nu)\frac{\partial^2 a_0}{\partial\theta^2}+(\lambda-3+4\nu)\frac{\partial b_0}{\partial\theta}=0$$
$$(\lambda-3+4\nu)\frac{\partial a_0}{\partial\theta}+(1-2\nu)(\lambda^2-1)b_0+2(1-\nu)\frac{\partial^2 b_0}{\partial\theta^2}=0. \quad (12)$$

Solving, by eliminating b_0 in favor of a_0 and its derivatives gives

$$\frac{\partial^4 a_0}{\partial\theta^4}+2(\lambda^2+1)\frac{\partial^2 a_0}{\partial\theta^2}+(\lambda^2-1)^2a_0=0,$$

and the roots of the corresponding auxiliary equation are

$$r=\pm i(\lambda\pm1).$$

The general solutions of equation (12) are therefore

$$a_0=K_0\cos(\lambda+1)\theta+L_0\sin(\lambda+1)\theta$$
$$+M_0\cos(\lambda-1)\theta+N_0\sin(\lambda-1)\theta$$
$$b_0=-K_0\sin(\lambda+1)\theta+L_0\cos(\lambda+1)\theta$$
$$+\frac{\lambda+3-4\nu}{\lambda-3+4\nu}[-M_0\sin(\lambda-1)\theta+N_0\cos(\lambda-1)\theta] \quad (13)$$

and the arbitrary constants K_0, L_0, M_0, N_0 are found from boundary conditions equation (10). These may be recast in terms of the zero-order coefficients of equation (11) to give

$$\frac{\partial b_0}{\partial\theta}+a_0+\frac{\nu}{1-\nu}\lambda a_0=0 \quad \text{when } \theta=\pm\pi$$
$$\frac{\partial a_0}{\partial\theta}+\lambda b_0-b_0=0. \quad (14)$$

Substituting equations (13) into (14) then gives the following relations between the arbitrary constants:

$$K_0\cos(\lambda+1)\pi\pm L_0\sin(\lambda+1)\pi$$
$$+\frac{\lambda-1}{\lambda-3+4\nu}[M_0\cos(\lambda-1)\pi\pm N_0\sin(\lambda-1)\pi]=0$$
$$\mp K_0\sin(\lambda+1)\pi+L_0\cos(\lambda+1)\pi$$
$$+\frac{\lambda-1}{\lambda-3+4\nu}[\mp M_0\sin(\lambda-1)\pi+N_0\cos(\lambda-1)\pi]=0. \quad (15)$$

The requirement that this set of equations has a nontrivial solution means that the determinant of the coefficients must vanish. This gives the following eigenvalues:

$$\lambda=\pm\frac{n}{2} \quad (n=0,\ 1,\ 2,\ \ldots) \quad (16)$$

and the requirement that the internal strain energy be bounded means that negative roots must be discarded. This result enables us to replace the double series (equation (11)) with the following single series

$$v_r=\sum_{n=0}^{\infty}r^{\frac{n}{2}}a_n(\theta)$$
$$v_\theta=\sum_{n=0}^{\infty}r^{\frac{n}{2}}b_n(\theta). \quad (17)$$

Solution

The intention now is to proceed in a similar way to that described in the previous section, which enabled λ to be found, in order to determine $a_n(\theta)$, $b_n(\theta)$. Thus, equations (17) are expanded and substituted into governing equations (9) and boundary conditions equation (10). Five terms were taken in our solution (including zero order), but the algebra for higher-order terms becomes extremely tedious. We shall therefore describe the solution for the first two terms in some detail, but present only the results for higher-order coefficients.

(a) Zero Order ($n=0$). The governing equations (9) become

$$-2(1-\nu)a_0+(1-2\nu)\frac{d^2a_0}{d\theta^2}-(3-4\nu)\frac{db_0}{d\theta}=0$$
$$(3-4\nu)\frac{da_0}{d\theta}-(1-2\nu)b_0+2(1-\nu)\frac{d^2b_0}{d\theta^2}=0. \quad (18)$$

Their general solution is

$$a_0=K_0\cos\theta+L_0\sin\theta+(3-4\nu)\theta(M_0\sin\theta-N_0\cos\theta)$$
$$b_0=-(K_0+M_0)\sin\theta+(L_0+N_0)\cos\theta$$
$$+(3-4\nu)\theta(M_0\cos\theta+N_0\sin\theta), \quad (19)$$

while the zero-order boundary conditions, from equations (10) are

$$\frac{db_0}{d\theta}+a_0=0 \quad \text{when } \theta=\pm\pi$$
$$\frac{da_0}{d\theta}-b_0=0. \quad (20)$$

Substitution of equations (19) into equations (20) reveals that $M_0=N_0=0$, so that the most general acceptable solution is

$$a_0=K_0\cos\theta+L_0\sin\theta$$
$$b_0=-K_0\sin\theta+L_0\cos\theta. \quad (21)$$

(b) First Order ($n=1$). The governing equations are

$$-\frac{3}{2}(1-\nu)a_1+(1-2\nu)\frac{d^2a_1}{d\theta^2}-\left(\frac{5}{2}-4\nu\right)\frac{db_1}{d\theta}=0$$
$$\left(\frac{7}{2}-4\nu\right)\frac{da_1}{d\theta}-\frac{3}{4}(1-2\nu)b_1+2(1-\nu)\frac{d^2b_1}{d\theta^2}=0 \quad (22)$$

whose solution is

$$a_1=K_1\cos\frac{3}{2}\theta+L_1\sin\frac{3}{2}\theta+M_1\cos\frac{\theta}{2}+N_1\sin\frac{\theta}{2}$$
$$b_1=-K_1\sin\frac{3}{2}\theta+L_1\cos\frac{3}{2}\theta-\frac{7-8\nu}{5-8\nu}\left(M_1\sin\frac{\theta}{2}-N_1\cos\frac{\theta}{2}\right). \quad (23)$$

The first-order boundary conditions are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\frac{db_1}{d\theta}+\frac{2-\nu}{2(1-\nu)}a_1=0 \quad \text{when } \theta=\pm\pi$$

$$\frac{da_1}{d\theta}-\frac{1}{2}b_1=0 \tag{24}$$

and invoking them gives the following relations

$$M_1=-(5-8\nu)K_1$$

$$N_1=-\frac{1}{3}(5-8\nu)L_1, \tag{25}$$

so that the most general acceptable first-order solution is

$$a_1=K_1\left[\cos\frac{3}{2}\theta-(5-8\nu)\cos\frac{\theta}{2}\right]+L_1\left[\sin\frac{3}{2}\theta-\frac{1}{3}(5-8\nu)\sin\frac{\theta}{2}\right]$$

$$b_1=-K_1\left[\sin\frac{3}{2}\theta-(7-8\nu)\sin\frac{\theta}{2}\right]+L_1\left[\cos\frac{3}{2}\theta-\frac{1}{3}(7-8\nu)\cos\frac{\theta}{2}\right]. \tag{26}$$

(c) Second and Higher Order ($n\geq 2$). The same general technique is followed as described in previous sections, and for brevity the algebra will be omitted here. Details are given in (Li and Hills, 1988) and the results are:

n = *2*

$$a_2=K_2(1-2\nu+\cos2\theta)-\nu(K_0\sin\theta_0+L_0\cos\theta_0)$$

$$b_2=-K_2\sin2\theta+N_2. \tag{27}$$

n = *3*

$$a_3=K_3\left[\cos\frac{5}{2}\theta+(3-8\nu)\cos\frac{\theta}{2}\right]+L_3\left[\sin\frac{5}{2}\theta+\frac{1}{5}(3-8\nu)\sin\frac{\theta}{2}\right]$$
$$+\frac{1}{4}(3-8\nu)\left[K_1\sin\left(\frac{3}{2}\theta+\theta_0\right)-\frac{1}{3}L_1\cos\left(\frac{3}{2}\theta+\theta_0\right)\right]$$
$$+\frac{1}{60}\left[(21+288\nu-384\nu^2)K_1\cos\theta_0\right.$$
$$\left.-\frac{1}{3}(107+96\nu-128\nu^2)L_1\sin\theta_0\right]\sin\frac{\theta}{2}$$
$$-\frac{1}{12}(13-96\nu+128\nu^2)\left(K_1\sin\theta_0-\frac{1}{3}L_1\cos\theta_0\right)\cos\frac{\theta}{2} \tag{28a}$$

$$b_3=-K_3\left[\sin\frac{5}{2}\theta-(9-8\nu)\sin\frac{\theta}{2}\right]+L_3\left[\cos\frac{5}{2}\theta-\frac{1}{5}(9-8\nu)\cos\frac{\theta}{2}\right]$$
$$+\frac{1}{4}(5-8\nu)\left[K_1\cos\left(\frac{3}{2}\theta+\theta_0\right)+\frac{1}{3}L_1\sin\left(\frac{3}{2}\theta+\theta_0\right)\right]$$
$$+\frac{1}{60}\left[(177-576\nu+384\nu^2)K_1\cos\theta_0\right.$$
$$\left.-\frac{1}{3}(79-192\nu+128\nu^2)L_1\sin\theta_0\right]\cos\frac{\theta}{2}$$
$$-\frac{1}{12}(55-192\nu+128\nu^2)(K_1\sin\theta_0-\frac{1}{3}L_1\cos\theta_0)\sin\frac{\theta}{2}. \tag{28b}$$

n = *4*

$$a_4=K_4\left[\cos3\theta+\frac{1}{3}(1-4\nu)\cos\theta\right]+L_4[\sin3\theta+(1-4\nu)\sin\theta]$$
$$+\left[\nu(1-\nu)K_2\cos\theta_0+\frac{1}{2}(1-\nu^2)(K_0\sin\theta_0+L_0\cos\theta_0)\cos\theta_0\right.$$
$$\left.+\frac{\nu}{2}N_2\sin\theta_0\right]\sin\theta+\left[-\frac{1}{3}(1-\nu)(2+\nu)K_2\sin\theta_0\right.$$
$$\left.+\frac{1}{6}(1+\nu)(2+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)\sin\theta_0-\frac{\nu}{2}N_2\cos\theta_0\right]\cos\theta \tag{29a}$$

$$b_4=-K_4\left[\sin3\theta-\frac{1}{3}(5-4\nu)\sin\theta\right]+L_4[\cos3\theta-(5-4\nu)\cos\theta]$$
$$-\left[\nu(1-\nu)K_2\cos\theta_0+\frac{1}{2}(1-\nu^2)(K_0\sin\theta_0+L_0\cos\theta_0)\cos\theta_0\right.$$
$$\left.+\frac{\nu}{2}N_2\sin\theta_0\right]\cos\theta+\left[\frac{1}{3}(1-\nu)(2-\nu)K_2\sin\theta_0\right.$$
$$\left.-\frac{1}{6}(1+\nu)(2-\nu)(K_0\sin\theta_0+L_0\cos\theta_0)\sin\theta_0-\frac{\nu}{2}N_2\cos\theta_0\right]\sin\theta. \tag{29b}$$

The solutions for $n=3$, 4 are worthy of comment, as the technique used involved some additional steps. First, the governing equations and boundary conditions are established, as before. Substitution of the boundary conditions generates inhomogeneous equations. These are solved by superposition of the solution to the corresponding homogeneous equation and a suitable particular integral.

Now that the displacement fields in the neighborhood of the crack tip are established, the corresponding stress state may be found by substitution into the geometric equations (4). Although the resulting expressions are useful they are also very lengthy, and they are recorded in the Appendix.

Finite Element Formulation

Previous sections have enabled the permitted variation in the stress and displacement fields near the crack tip, in a set of coordinates natural to the given problem (Fig. 1), to be found. These may be used to establish a special finite element incorporating these characteristics, and which may be used within a mesh of ordinary, eight-noded isoparametric elements (Fig. 2) to permit good accuracy to be achieved at the expense of little speed.

The element itself is readily established in the usual way. The relationship between the stress state and the coefficients of the expansion are given in the Appendix, and may be used to develop a stress matrix [*P*] defined by

$$\{\sigma\}=\mathbf{P}\{\alpha\} \tag{30}$$

where

$$\{\sigma\}=(\sigma_{\rho\rho},\ \sigma_{\theta\theta},\ \sigma_{\phi\phi},\ \sigma_{\rho\theta})^T$$
$$\{\alpha\}=(K_0,\ L_0,\ K_1,\ L_1,\ \ldots,\ K_4,\ L_4)^T.$$

Also, the stress and strain components are related by Hooke's law, i.e.,

$$\{\epsilon\}=\mathbf{C}\{\sigma\}=\mathbf{CP}\{\alpha\} \tag{31}$$

where $\{\epsilon\}=(\epsilon_{\rho\rho},\ \epsilon_{\theta\theta},\ \epsilon_{\phi\phi},\ \epsilon_{\rho\theta})^T$ and

$$\mathbf{C}=\frac{1}{E}\begin{vmatrix} 1 & -\nu & -\nu & \\ -\nu & 1 & -\nu & \\ -\nu & -\nu & 1 & \\ & & & 2(1+\nu) \end{vmatrix}.$$

Therefore, the stiffness matrix K_α corresponding to $\{\alpha\}$ may be defined by

$$\mathbf{K}_\alpha=\int_s\mathbf{P}^T\mathbf{CP}ds. \tag{32}$$

The connection between the special element and the ordinary elements is expressed by ensuring that displacements at shared nodes are equal, i.e.,

$$\mathbf{A}\{\alpha\}=\{u\} \tag{33}$$

where $\{u\}=(u_{\rho i},\ u_{\theta i})^T$, $i=1,\ 2,\ \ldots,\ n$, n being the number of common nodes. $\mathbf{A}\{\alpha\}$ denotes the nodal displacements of the special element obtained from the coefficients of the expansion. The matrix $\mathbf{A}$ can be found from equations (17) where $a_n(\theta)$, $b_n(\theta)$ are given by equations (21), (26), (27), (28), and

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

(29). However, since the dimensions of $\{u\}$ and $\{\alpha\}$ are not usually the same, a least-squares fit has to be performed. To implement, this equation (33) may be rewritten in the form

$$\mathbf{A}^T\mathbf{A}\{\alpha\} = \mathbf{A}^T\{u\}$$

i.e.,

$$\mathbf{N}\{\alpha\} = \mathbf{A}^T\{u\}$$

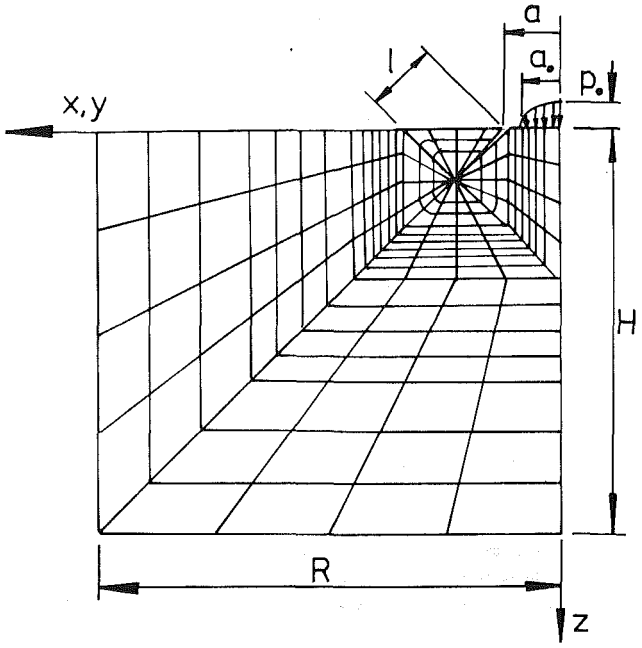


Fig. 2 Finite element mesh

where $\mathbf{N} = \mathbf{A}^T\mathbf{A}$. Hence,

$$\{\alpha\} = \mathbf{N}^{-1}\mathbf{A}^T\{u\} = \mathbf{T}\{u\} \tag{34}$$

where $\mathbf{T} = \mathbf{N}^{-1}\mathbf{A}^T$, a transformation matrix.

Thus, the stiffness matrix corresponding to $\{u\}$ can be derived in the form

$$\mathbf{K} = \mathbf{T}^T\mathbf{K}_\alpha\mathbf{T}. \tag{35}$$

An advantage of using the type of special element described above is that the crack-tip stress intensity factors may be abstracted immediately. The definition of stress intensity factors is

$$K_I = \lim_{\rho\to 0}\sqrt{2\pi\rho}\,\sigma_{\theta|\theta=0} \tag{36a}$$

$$K_{II} = \lim_{\rho\to 0}\sqrt{2\pi\rho}\,\sigma_{\rho\theta|\theta=0} \tag{36b}$$

and it will be recognized that one of the terms in the eigenstrain expansion corresponds to a $1/\sqrt{\rho}$ variation (see equation (A2)). Thus,

$$K_I = -\sqrt{2\pi(a+l\sin\theta_0)}\,\frac{2EK_1}{1+\nu} \tag{37a}$$

$$K_{II} = \sqrt{2\pi(a+l\sin\theta_0)}\,\frac{2EL_1}{3(1+\nu)}. \tag{37b}$$

Results

In order to check the manipulation and validate the code, we re-evaluated a simple problem with an exact solution. This was a three-dimensional space containing a penny crack under uniform tension so that $a=0$, $\theta_0=\pi/2$, and the analytical solution to which can be expressed as follows:

$$K_I = \frac{2}{\pi}\sigma\sqrt{\pi l} \tag{38}$$

where σ is the tension and l is the crack radius. When $\sigma=2$, $l=1$, the analytical value of K_I is 2.257 from equation (38),

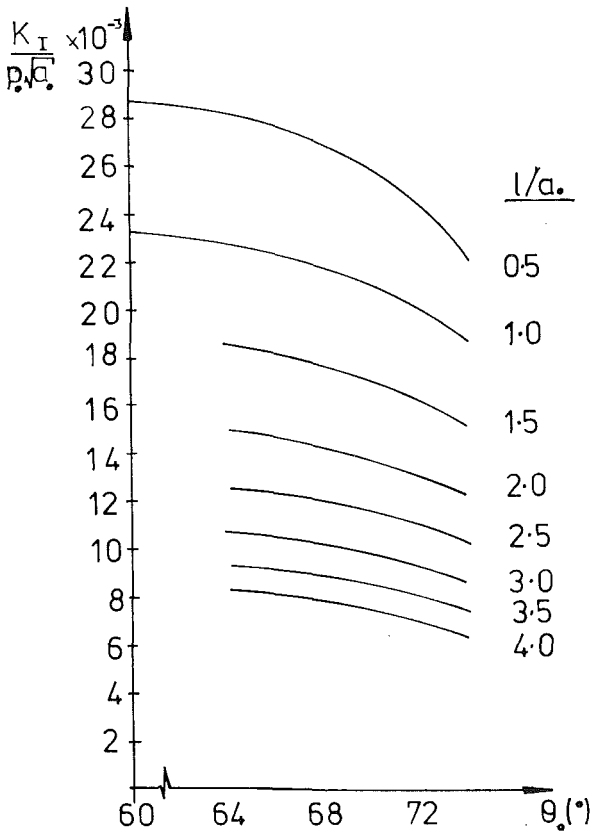


Fig. 3 Dimensionless stress intensity factors, k_I

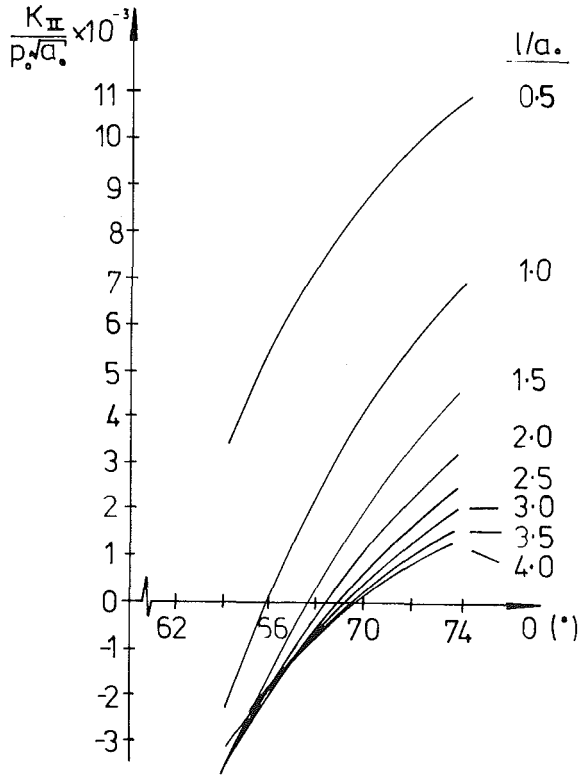


Fig. 4 Dimensionless stress intensity factors, k_{II}

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1(*a*) Dimensionless stress intensity factors $k_I = \dfrac{K_I}{p_0\sqrt{a_0}}$ for the Hertzian cone crack

$$\left(\frac{a}{a_0} = 1.05, \quad E = 68950 \text{ MPa}, \quad \nu = 0.23\right)$$

$\frac{l}{a_0}$	θ_0									
	56 deg	58 deg	60 deg	62 deg	64 deg	66 deg	68 deg	70 deg	72 deg	74 deg
0.5	.0292	.0294	.0287	.0284	.0278	.0271	.0262	.0251	.0238	.0224
1.0	.0232	.0233	.0233	.0231	.0228	.0223	.0217	.0209	.0200	.0189
1.5	.0190	.0190	.0189	.0186	.0183	.0179	.0174	.0168	.0161	.0153
2.0	.0160	.0159	.0157	.0154	.0151	.0147	.0143	.0138	.0132	.0125
2.5	.0137	.0135	.0133	.0130	.0127	.0123	.0120	.0115	.0110	.0104
3.0	.0119	.0117	.0115	.0112	.0109	.0105	.0102	.0098	.0094	.0089
3.5	.0105	.0103	.0100	.0097	.0095	.0092	.0088	.0085	.0081	.0077
4.0	.0094	.0091	.0089	.0086	.0083	.0080	.0078	.0074	.0071	.0067
4.5	.0084	.0082	.0079	.0077	.0074	.0071	.0069	.0066	.0063	.0059
5.0	.0076	.0074	.0071	.0069	.0067	.0064	.0062	.0059	.0056	.0053

Table 1(*b*) Dimensionless stress intensity factors $k_{II} = \dfrac{K_{II}}{p_0\sqrt{a_0}}$ for the Hertzian cone crack

$$\left(\frac{a}{a_0} = 1.05, \quad E = 68950 \text{ MPa}, \quad \nu = 0.23\right)$$

$\frac{l}{a_0}$	θ_0									
	56 deg	58 deg	60 deg	62 deg	64 deg	66 deg	68 deg	70 deg	72 deg	74 deg
0.5	−.0082	−.0049	−.0018	.0009	.0033	.0054	.0072	.0088	.0100	.0108
1.0	−.0143	−.0109	−.0077	−.0048	−.0022	.0002	.0023	.0041	.0056	.0068
1.5	−.0144	−.0114	−.0086	−.0060	−.0036	−.0015	.0004	.0020	.0034	.0046
2.0	−.0130	−.0104	−.0081	−.0059	−.0039	−.0020	−.0004	.0010	.0023	.0033
2.5	−.0114	−.0093	−.0072	−.0054	−.0037	−.0021	−.0007	.0006	.0016	.0025
3.0	−.0100	−.0081	−.0064	−.0048	−.0033	−.0020	−.0008	.0003	.0012	.0020
3.5	−.0088	−.0071	−.0056	−.0043	−.0030	−.0018	−.0008	.0002	.0010	.0017
4.0	−.0077	−.0063	−.0050	−.0038	−.0027	−.0016	−.0007	.0001	.0008	.0014
4.5	−.0068	−.0056	−.0044	−.0034	−.0024	−.0015	−.0007	.0001	.0007	.0013
5.0	−.0061	−.0050	−.0040	−.0030	−.0021	−.0013	−.0006	.0001	.0006	.0011

while our numerical value is $K_I = 2.265$. The error is equal to 0.35 percent.

The closeness of the agreement means that this partial check on our coding was successful, and improves our confidence in results obtained from a full three-dimensional analysis.

A second check on overall accuracy was made by determining the contact stress due to a Hertzian pressure distribution in the crack's absence, but using essentially the same mesh (without the special crack-tip elements). It is notoriously difficult to obtain good representations of contact problems, and we found it necessary to set $R = H = 25a$ (Fig. 2), to attain satisfactory results. With these overall dimensions the stresses within the neighborhood of the contact were within 0.5 percent of the exact solution (Hüber, 1904).

Although recent work has indicated the importance of interfacial shear tractions between the ball and specimen in deciding the precise location of *initiation* of the cone crack (Johnson et al., 1973), their influence on the stress state is very short range since they are self-equilibrating. A full analysis of the internal stress field due to a frictional contact has been carried out (Hills and Sackfield, 1987), and beyond a depth of about $z = 0.5a$ the disturbance to the frictionless (Hertzian) stress field is quite negligible. Thus, the influence of interfacial shear on a propagating crack will also be trivial, and we feel justified in omitting it from the present calculation. Further, it is known that the influence of surface or near-surface defects on contact pressure distribution is small (Bryant et al., 1984), and we have therefore used type 1 (imposed traction) loading with

$$p(r) = p_0\sqrt{1 - (r/a_0)^2} \tag{39}$$

where r is the radial coordinate, a_0 the radius of contact, and p_0, the peak contact pressure, is given by

$$p_0 = \frac{3P}{2\pi a_0^2}, \tag{40}$$

P being the applied load.

The crack is assumed to initiate at a radius $a (a \geq a_0)$, Fig. 2, which also depicts the mesh used. The value of a/a_0 used was 1.05, which is the one found in practice (Johnson et al., 1973). Values of the stress intensity factors found are shown in Tables 1(*a*) and 1(*b*) with elastic constants appropriate for glass and in Figs. 3 and 4 for appropriate crack lengths l and angles θ_0, where θ_0 denotes the angle between the crack and vertical axis z. It should be recognized that, except at a specific angle (which is a function of crack length, Fig. 4), the crack tip suffers a shear loading, which may change sign as the crack grows. Thus, it might be expected that the cone crack is propelled by combined-mode loading, although the crack faces seem to be almost straight when viewed in the r—z plane.

It may be seen that the normalized stress intensity factors decrease with increasing length rapidly. From $\theta_0 = 74$ deg to $\theta_0 = 60$ deg the variation of K_I with θ_0 is fairly slow, and K_{II} is about 10 percent of K_I. It would seem that the K_I reaches the maximum value when θ_0 is about 60 deg, but it is sensitive to the value of Poisson's ratio and the crack length. If θ_0 decreases continuously, both K_I and K_{II} decrease and become negative, so that the crack tip closes.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The use of the calibration factors presented in determining fracture toughness values by means of inducing cone cracks in described elsewhere (Hills, Li and Wang, 1989). Values found for glass by this technique compare very closely with those obtained by bulk testing techniques.

Discussion and Conclusion

In this paper a so-called "Local-global analysis" has been used to deal with an axisymmetric fracture problem. A general solution of the stress-strain fields at the crack tip, which suffers both mode I and mode II and whose geometry is that of a half space containing a cone crack, has been found. In an analogous scheme to the Williams expansion technique for a plane fracture problem, the mechanical behavior near the crack tip was revealed and provides a better foundation for numerical fracture analyses.

Based on the stress-strain fields obtained, a high-order special element was established to substitute the dense mesh otherwise needed near the crack tip.

The stress intensity factors for a half space containing a cone crack were then calculated, and the variation of stress intensity factors with geometric parameters is investigated. They can be used to determine fracture toughness for brittle materials, such as glass and ceramics.

The crack angle, θ_0, found in tests varies slightly, according to test conditions. In his classic experiments, Roesler (1957) found values of 68.5 deg ± 1 deg. However, these were performed using a near flat-ended indenter, so that the contact patch size remains constant as the load is applied. In our own tests (Li and Hills, 1988) utilizing a steel ball there was a somewhat wider variation in cone semi-angle, and we obtained values in the range 56 deg $< \theta_0 <$ 72 deg. It may be seen that these values lie in the region where $K_{II} \approx 0$, Fig. 4 and indeed for $l/a_0 \approx 3$, $\theta_0 = 69$ deg, which is where the bulk of our experimental data lies, K_{II} does indeed vanish.

References

Bryant, M. D., Miller, G. R., and Keer, L. M., 1984, "Line contact between a rigid indenter and a damaged elastic body," *Q. J. Appl. Math.*, Vol. 37, No. 3, pp. 467-478.

Finnie, I., Dolev, D., and Khatibloo, M., 1981, "On the physical basis of Auerbach's law," *J. of Engineering Materials and Technology*, Vol. 103, pp. 183-184.

Frank, F. C., and Lawn, B. R., 1967, "On the theory of Hertzian fracture," *Proc. Roy. Soc.*, Vol. 299, pp. 291-306.

Hills, D. A., and Sackfield, A., 1987, "The stress field induced by normal contact between dissimilar spheres," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 8-14.

Hills, D. A., Li Yingzhi, and Wang, Z. P., 1989, "Measurement of the fracture toughness of brittle materials by indentation," *Proceedings of the third conference on solid mechanics*, I. M. Allison and C. Ruiz, eds., Elsevier, New York, pp. 55-65.

Hüber, M. T., 1904, "Zur theorie der Beruhrung fester elastische Korper," *Ann. der Phys.*, Vol. 14, pp. 153.

Johnson, K. L., O'Connor, J. J., and Woodward, A. C., 1973, "The effect of the indenter elasticity on the Hertzian fracture of brittle materials," *Proc. Roy. Soc.*, Vol. 334, pp. 95-117.

Li Yingzhi, 1988, "Fracture analyses for three-dimensional bodies with surface crack," *Proceedings of Symposium on Surface Crack Growth: Models, Experiments, Structures*, Reno, Nev., ASTM-STP 1060, in press.

Li Yingzhi, and Hills, D. A., 1988, "The Hertzian cone crack: theoretical, numerical and experimental studies," Research Report No. OUEL 1743/88, Department of Engineering Science, Oxford University, U. K.

Liu Chuntu, and Li Yingzhi, 1987, "Study on cracked plates, shells and three-dimensional bodies," *Engineering Fracture Mechanics* (a special issue in honor of Prof. Yokobori on the occasion of his 70 birthday), Vol. 28.

Roesler, F. C., 1957, "Brittle fractures near equilibrium," *Proc. Phys. Soc.*, Vol. 69, pp. 981-992.

Williams, M. L., 1957, "On the stress distribution at the base of a stationary crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, pp. 109-114.

Yahsi, O. S., and Parnas, L., 1986, "Conical crack problem in semi-infinite media with stress-free boundary condition," *Int. J. of Fracture*, Vol. 31, pp. 291-301.

Zak, A. R., 1972, "Elastic analysis of cylindrical configurations with stress singularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 501-506.

APPENDIX

Stress Fields at the Crack Tip

(*a*) **Zero Order** ($n = 0$).

$$\sigma_{\rho 0} = \sigma_{\theta 0} = \sigma_{\phi 0} = \sigma_{\rho\theta 0} = 0 \tag{A1}$$

(*b*) **First Order** ($n = 1$).

$$r^{\frac{1}{2}}\sigma_{\rho 1} = \frac{E}{2(1+\nu)}\left[K_1\left(\cos\frac{3}{2}\theta - 5\cos\frac{\theta}{2}\right) + L_1\left(\sin\frac{3}{2}\theta - \frac{5}{3}\sin\frac{\theta}{2}\right)\right]$$

$$r^{\frac{1}{2}}\sigma_{\theta 1} = -\frac{E}{2(1+\nu)}\left[K_1\left(\cos\frac{3}{2}\theta + 3\cos\frac{\theta}{2}\right) + L_1\left(\sin\frac{3}{2}\theta + \sin\frac{\theta}{2}\right)\right]$$

$$r^{\frac{1}{2}}\sigma_{\phi 1} = -\frac{4E\nu}{1+\nu}\left(K_1\cos\frac{\theta}{2} + \frac{1}{3}L_1\sin\frac{\theta}{2}\right)$$

$$r^{\frac{1}{2}}\sigma_{\rho\theta 1} = \frac{E}{2(1+\nu)}\left[-K_1\left(\sin\frac{3}{2}\theta + \sin\frac{\theta}{2}\right) + L_1\left(\cos\frac{3}{2}\theta + \frac{1}{3}\cos\frac{\theta}{2}\right)\right] \tag{A2}$$

(*c*) **Second Order** ($n = 2$).

$$\sigma_{\rho 2} = \frac{E}{1+\nu}K_2(1+\cos 2\theta)$$

$$\sigma_{\theta 2} = \frac{E}{1+\nu}K_2(1-\cos 2\theta)$$

$$\sigma_{\phi 2} = \frac{2E\nu}{1+\nu}K_2 + E(K_0\sin\theta_0 + L_0\cos\theta_0)$$

$$\sigma_{\rho\theta 2} = -\frac{E}{1+\nu}K_2\sin 2\theta \tag{A3}$$

(*d*) **Third Order** ($n = 3$).

$$r^{-\frac{1}{2}}\sigma_{\rho 3} = \frac{3E}{2(1+\nu)}\left[K_3\left(\cos\frac{5}{2}\theta + 3\cos\frac{\theta}{2}\right) + L_3\left(\sin\frac{5}{2}\theta + \frac{3}{5}\sin\frac{\theta}{2}\right)\right]$$
$$+\frac{E}{1+\nu}\left\{\frac{1}{8}(9-16\nu)\left[K_1\sin\left(\frac{3}{2}\theta+\theta_0\right) - \frac{1}{3}L_1\cos\left(\frac{3}{2}\theta+\theta_0\right)\right]\right.$$
$$+\frac{1}{40}\left[(21-16\nu)K_1\cos\theta_0 - \frac{1}{3}(107-112\nu)L_1\sin\theta_0\right]\sin\frac{\theta}{2}$$
$$\left.+\frac{1}{24}\left[-(39-48\nu)K_1\sin\theta_0 + (13-80\nu)L_1\cos\theta_0\right]\cos\frac{\theta}{2}\right\}$$

$$r^{-\frac{1}{2}}\sigma_{\theta 3} = \frac{3E}{2(1+\nu)}\left[K_3\left(-\cos\frac{5}{2}\theta + 5\cos\frac{\theta}{2}\right) + L_3\left(-\sin\frac{5}{2}\theta + \sin\frac{\theta}{2}\right)\right]$$
$$+\frac{E}{8(1+\nu)}\left\{-(9-16\nu)\left[K_1\sin\left(\frac{3}{2}\theta+\theta_0\right) - \frac{1}{3}L_1\cos\left(\frac{3}{2}\theta+\theta_0\right)\right]\right.$$
$$-(9-16\nu)\left(K_1\cos\theta_0 + \frac{1}{3}L_1\sin\theta_0\right)\sin\frac{\theta}{2}$$
$$\left.+\left[-(27-48\nu)K_1\sin\theta_0 + \frac{1}{3}(27-112\nu)L_1\cos\theta_0\right]\cos\frac{\theta}{2}\right\}$$

$$r^{-\frac{1}{2}}\sigma_{\phi 3} = \frac{12E\nu}{1+\nu}\left(K_3\cos\frac{\theta}{2} + \frac{1}{5}L_3\sin\frac{\theta}{2}\right)$$
$$+E\left[K_1\sin\left(\frac{3}{2}\theta+\theta_0\right) - \frac{1}{3}L_1\cos\left(\frac{3}{2}\theta+\theta_0\right)\right]$$
$$+\frac{E}{1+\nu}\left\{\frac{1}{5}\left[(-35+2\nu+48\nu^2)K_1\cos\theta_0\right.\right.$$
$$\left.\left.+\frac{1}{3}(45-14\nu-16\nu^2)L_1\sin\theta_0\right]\sin\frac{\theta}{2}\right.$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$-\left[(5+2\nu-16\nu^2)K_1\sin\theta_0+\frac{1}{3}(3-10\nu+16\nu^2)L_1\cos\theta_0\right]\cos\frac{\theta}{2}\Bigg\}$$

$$r^{-\frac{1}{2}}\sigma_{\rho\theta 3}=\frac{3E}{2(1+\nu)}\left[K_3\left(-\sin\frac{5}{2}\theta+\sin\frac{\theta}{2}\right)+L_3\left(\cos\frac{5}{2}\theta-\frac{1}{5}\cos\frac{\theta}{2}\right)\right]$$
$$+\frac{E}{8(1+\nu)}\Bigg\{(7-16\nu)\left[K_1\cos\left(\frac{3}{2}\theta+\theta_0\right)+\frac{1}{3}L_1\sin\left(\frac{3}{2}\theta+\theta_0\right)\right]$$
$$+\frac{1}{5}\left[(33-48\nu)K_1\cos\theta_0+\frac{1}{3}(-31+16\nu)L_1\sin\theta_0\right]\cos\frac{\theta}{2}$$
$$-(7-16\nu)\left(K_1\sin\theta_0-\frac{1}{3}L_1\cos\theta_0\right)\sin\frac{\theta}{2}\Bigg\} \qquad \text{(A4)}$$

(e) Fourth Order ($n=4$).

$$r^{-1}\sigma_{\rho 4}=\frac{E}{1+\nu}\Bigg\{2K_4\left(\cos3\theta+\frac{1}{3}\cos\theta\right)+2L_4(\sin3\theta+\sin\theta)$$
$$+[2\nu K_2+(1+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)]\cos\theta_0\sin\theta$$
$$+\left[-\frac{4}{3}(1-\nu)K_2+\frac{2}{3}(1+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)\right]\sin\theta_0\cos\theta\Bigg\}$$

$$r^{-1}\sigma_{\theta 4}=-\frac{2E}{1+\nu}[K_4(\cos3\theta-\cos\theta)+L_4(\sin3\theta-3\sin\theta)]$$
$$+\frac{E}{1+\nu}[2\nu K_2+(1+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)]\cos\theta_0\sin\theta$$

$$r^{-1}\sigma_{\phi 4}=\frac{E}{1+\nu}\Bigg\{\frac{8}{3}\nu K_4\cos\theta+8\nu L_4\sin\theta$$
$$-[2\nu(1-\nu)K_2\cos\theta_0+(1-\nu^2)(K_0\sin\theta_0+L_0\cos\theta_0)\cos\theta_0$$
$$+(1+\nu)N_2\sin\theta_0]\sin\theta+\left[\frac{2}{3}(3+\nu)(1-\nu)K_2\sin\theta_0\right.$$
$$\left.-\frac{1}{3}(3+\nu)(1+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)\cos\theta_0+(1+\nu)N_2\cos\theta_0\right]\cos\theta\Bigg\}$$

$$r^{-1}\sigma_{\rho\theta 4}=\frac{E}{1+\nu}\Bigg\{-2K_4\left(\sin3\theta-\frac{1}{3}\sin\theta\right)+2L_4(\cos3\theta-\cos\theta)$$
$$+\left[\frac{2}{3}(1-\nu)K_2\sin\theta_0-\frac{1}{3}(1+\nu)(K_0\sin\theta_0+L_0\cos\theta_0)\sin\theta_0\right]\sin\theta\Bigg\} \qquad \text{(A5)}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Rigid Punch on Stressed Orthotropic Half-Plane With Partial Slip

H. L. Bjarnehed
Research Associate,
Division of Solid Mechanics,
Chalmers University of Technology,
S-412 96 Göteborg,
Sweden

An orthotropic half-plane carrying a transverse uniaxial stress and loaded perpendicularly on the free edge by a rigid flat punch is considered. Governing integral equations are stated for the unknown contact stresses between the punch and the half-plane. An analytical solution for contact with complete adhesion and a numerical solution with limiting friction, according to the Coulomb friction law, are presented. Besides distributions of contact stresses the numerical results show relations between vertical load, applied uniform foundation stress, coefficient of Coulomb friction, and size of the slip region.

1 Introduction

The background of the title problem is a steel clamping joint attached in the middle of a composite truck leaf spring. Cracks and delamination have been observed in the clamped region and wear has been detected in the contact area of such springs used under running conditions. It is of main interest to know how interface stresses and slip zones will be affected by bending in the composite leaf spring.

In this study the contact is modeled as a rigid flat punch (clamper) indented vertically into a uniformly stressed orthotropic half-plane (representing bending stress at the surface of the composite leaf spring) as shown in Fig. 1. The analysis is formulated with the assumption of a single parameter of loading, defined as the ratio of the applied uniform foundation stress, and the mean value of the interface pressure (proportional loading).

The problem of a rigid flat punch on an isotropic half-plane with partial slip was analyzed by Spence (1973). He considered a punch which is loaded vertically and assumed that Coulomb friction applies over the entire edge of contact. He divided the contact region into an inner part with adhesion, friction being sufficient to prevent slip in a central region, surrounded by two regions of slip. Spence formulated a single homogeneous Fredholm equation whose largest eigenvalue gives the extent of the adhesive region in terms of Poisson's ratio ν and the coefficient of Coulomb friction μ. Guided by this eigenvalue solution he numerically solved coupled integral equations for the unknown interface stresses. The related problem of vertical indentation of an isotropic half-plane by a rigid flat punch, with given regions of adhesion and Coulomb friction, was studied by Sundelius (1981). Among other things he studied the case of a single centrally placed region of adhesion, the dimension of which is known beforehand, surrounded with two regions with Coulomb friction. For different combinations of the values of Poisson's ratio ν and coefficient of friction μ (including the eigenvalue solution derived by Spence) he calculated distributions of the interface stresses.

Also in the present study the Coulomb friction law is applied in the contact zone. Within an interval of loading it is found that slip zones appear at the outer ends of the contact zone as described above (Spence). The treatment of this problem shows that the extent of the adhesive region depends not only on material parameters and coefficient of Coulomb friction but also on this single parameter of loading. Outside this interval of loading more than two slip zones may appear in the contact zone. Coupled integral equations for the unknown contact stresses are found, similar to those found by Spence. These equations are solved numerically in the same way as was done by Spence, but with given values of the extent of the adhesive region, material parameters, and coefficient of Coulomb friction this solution also gives the proper value of the single parameter of loading. In order to formulate the contact problem with finite friction, it is necessary to study the interface stresses for the case of complete adhesion over the entire edge

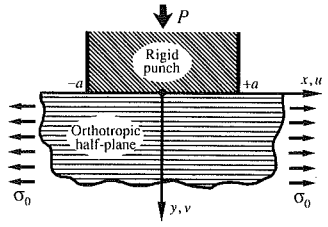


Fig. 1 Vertical indentation by a rigid flat punch into the edge of an stressed orthotropic half-plane

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16–19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 2, 1990; final revision, Apr. 10, 1990.

Paper No. 91-APM-11.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of contact. Therefore, an analytical solution is derived for the case of complete adhesion between the punch and the half-plane. The difference in the governing integral equations for the unknown interface stresses and their solution, compared to the corresponding indentation problem with an isotropic half-plane, is one of different constants. An analysis of the interface pressure for frictionless contact shows it to be identical to the isotropic solution (frictionless contact is not treated in this study).

2 Problem Statement

Consider the orthotropic homogeneous elastic half-plane in Fig. 1. The principal directions of elasticity are parallel and perpendicular to the horizontal edge. The half-plane is indented over the finite part of the edge $|x| \le a$ by a vertically loaded (P) rigid flat punch. In addition there is a load of uniform stress (σ_0) in the half-plane applied in the x-direction.

The friction between the two bodies in contact is assumed to be limited. This means that slip occurs in parts of the contact region. The Coulomb friction law will be used in this study.

3 Governing Equations

The main equations relating the surface stresses and displacements in the region $|x| \le a$ can be derived either by using complex analysis (see Muskhelishvili (1949)) or by means of Fourier transforms, which is of advantage in the case of an orthotropic half-plane as shown in the Appendix. These equations read (with $\int$ denoting the principal value of the integral)

$$\int_{-1}^{1} \frac{\tau(S)}{X-S}\, dS = -C_1\sigma(X) + C_2 u'(X) - C_3\sigma_0, \quad |X| \le 1, \quad (1a)$$

$$\int_{-1}^{1} \frac{\sigma(S)}{X-S}\, dS = C_4\tau(X) + C_5 v'(X), \quad |X| \le 1. \quad (1b)$$

These are coupled singular integral equations of the second kind, where u and v denote the components of surface displacements in the directions x and y, respectively (Fig. 1). The components of surface shear and normal stresses, i.e., τ and σ (Fig. 7), and the material parameters C_1, C_2, C_3, C_4, and C_5 (equations 26) are defined in the Appendix, and

$$X = \frac{x}{a}, \quad S = \frac{s}{a}, \quad u'(X) = \frac{1}{a}\frac{du}{dX}, \quad v'(X) = \frac{1}{a}\frac{dv}{dX}. \quad (2a,b,c,d)$$

4 Complete Adhesion

By aid of the no-slip solution of equations (1), it is possible to determine the location of the slip zones (see Section 5). Therefore, assume that the friction at the interface is sufficient to prevent any slip between the punch and the edge of the half-plane. The two surfaces in contact can be considered as glued together. We seek the distribution of both normal (σ) and shear (τ) stresses within the contact region. With the boundary conditions within the contact zone, i.e., $Y=0$ and $|X| \le 1$,

$$u'(X) = 0, \quad v'(X) = 0, \quad (3a,b)$$

equations (1) give

$$\int_{-1}^{1} \frac{\tau(S)}{X-S}\, dS = -C_1\sigma(X) - C_3\sigma_0, \quad \int_{-1}^{1} \frac{\sigma(S)}{X-S}\, dS = C_4\tau(X), \quad |X| \le 1, \quad (4a,b)$$

which are coupled singular integral equations of the second kind for the unknown interface stresses σ and τ. The general solution of coupled singular integral equations of the second kind is given by Söhngen (1954) (see the Appendix). The solutions of equations (4) read

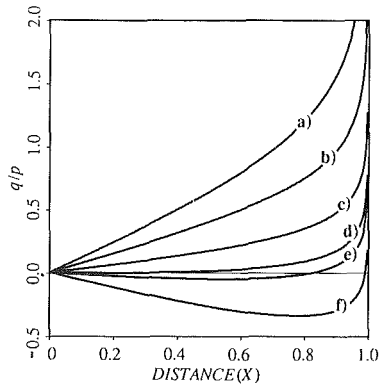


Fig. 2 Distribution of the ratio q/p between shear stress q and pressure p, within the region of contact $0 \le X \le 1$, for the case of complete adhesion: a) $\tilde{\sigma}_0 = -2$, b) $\tilde{\sigma}_0 = -1$, c) $\tilde{\sigma}_0 = 0$, d) $\tilde{\sigma}_0 = \tilde{\sigma}^* \approx 0.7$, e) $\tilde{\sigma}_0 = 1$, f) $\tilde{\sigma}_0 = 2$. (Material properties—see Section 6.)

$$p(X) = \frac{\cosh(\pi\eta)}{\sqrt{1-X^2}} \left\{ \frac{2}{\pi} \cos\psi(X) + \tilde{\sigma}_0\kappa \tanh(\pi\eta)[X \sin\psi(X) - 2\eta \cos\psi(X)] \right\}, \quad |X| \le 1, \quad (5a)$$

$$q(X) = \rho\frac{\cosh(\pi\eta)}{\sqrt{1-X^2}} \left\{ \frac{2}{\pi} \sin\psi(X) - \tilde{\sigma}_0\kappa \tanh(\pi\eta)[X \cos\psi(X) + 2\eta \sin\psi(X)] \right\}, \quad |X| \le 1, \quad (5b)$$

where

$$p(X), q(X) = -\frac{\sigma(X), \tau(X)}{(P/2a)}, \quad \psi(X) = \eta \ln\left(\frac{1+X}{1-X}\right), \quad (6a,b,c)$$

$$\tilde{\sigma}_0 = \frac{\sigma_0}{(P/2a)}, \quad \kappa = \frac{C_3}{C_1}, \quad (6d,e)$$

$$\rho = \sqrt{\frac{C_1}{C_4}}, \quad \eta = \frac{1}{2\pi} \ln\left(\frac{\pi + \sqrt{C_1C_4}}{\pi - \sqrt{C_1C_4}}\right). \quad (6f,g)$$

5 Partial Slip

Now we assume finite friction, with the coefficient of Coulomb friction, μ, between the edges in contact. Hence, $|\tau| \le \mu|\sigma|$ within the contact region with equality only if the friction is not sufficient to prevent slip. A study of the distribution of the ratio q/p in the case of complete adhesion, for different values of $\tilde{\sigma}_0$, gives a hint as to where slip can be expected—because $|q/p|$ indicates the necessary coefficient of friction to prevent slip. Since the distribution of q/p is antisymmetric with respect to the Y-axis, it is only necessary to study half the interval $|X| \le 1$; for example, $0 \le X \le 1$ as shown in Fig. 2. When $\tilde{\sigma}_0 \le \tilde{\sigma}^*$, where the value of $\tilde{\sigma}^*$ is positive and read as follows, $4\eta/(1+4\eta^2)/(\pi\kappa\tanh(\pi\eta))$, the ratio q/p is monotonically increasing and approaching infinite values at the edge $X=1$. With limiting friction, the conclusion is that slip zones occur at the outer ends of the contact zone. Further, when $\tilde{\sigma}_0 > \tilde{\sigma}^*$, Fig. 2 shows a possibility of more than two slip zones. This is due to the fact that the ratio $|q/p|$ has a maximum point, besides the singular point at $X=1$, in the region $X<1$. However, we will assume that slip only exists in the region $C < |X| \le 1$, and consequently, there is complete adhesion in the region $|X| \le C$. But for every combination of material parameters and coefficient of friction, there will be needed an investigation of the validity of the solution, i.e., a determination of the permitted range of $\tilde{\sigma}_0$ Thus, $|\tau(X)| \le \mu|\sigma(X)|$ in the adhesive region $|X| \le C$ with equality only at $|X| = C$ and $|\tau(X)| = \mu|\sigma(X)|$ in the slip zones $C < |X| \le 1$.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

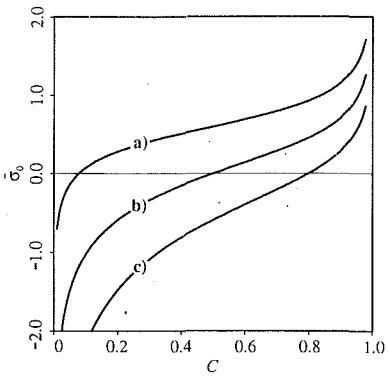


Fig. 3 Dependence of size of adhesive region on loading parameter: a) $\mu = 0.2$, b) $\mu = 0.4$, c) $\mu = 0.6$

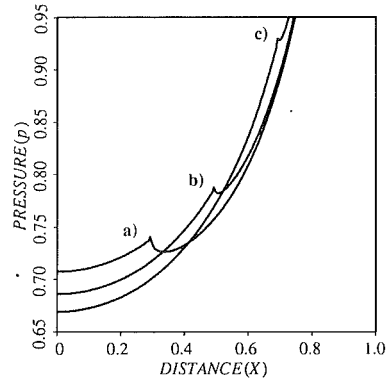


Fig. 4 Pressure distributions within the interval $0 \le X \le 1$ ($\mu = 0.4$): a) $\tilde{\sigma}_0 = -0.33$ $C = 0.30$, b) $\tilde{\sigma}_0 \approx 0$ $C = 0.50$, c) $\tilde{\sigma}_0 = 0.29$ $C = 0.70$

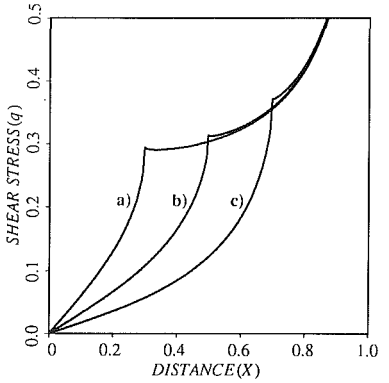


Fig. 5 Shear stress distributions within the interval $0 \le X \le 1$ ($\mu = 0.4$): a) $\tilde{\sigma}_0 = -0.33$ $C = 0.30$, b) $\tilde{\sigma}_0 \approx 0$ $C = 0.50$, c) $\tilde{\sigma}_0 = 0.29$ $C = 0.70$

During loading we assume that the ratio $\sigma_0/(P/2a)$ $(=\tilde{\sigma}_0)$, where $P > 0$, is constant, and that the loads P and σ_0 are applied monotonically (proportional loading) so that slip always occurs in the same direction (inwards toward the center $X = 0$). Furthermore, we assume that the loads are applied sufficiently slowly for static equilibrium to apply at each stage.

With given values of C, μ, and material parameters we seek the distributions of both normal (σ) and shear (τ) stresses within the region of contact. In addition, the solution will give the proper value of $\tilde{\sigma}_0$, because there is just one value of $\tilde{\sigma}_0$ for which the inequality $|\tau| \le \mu|\sigma|$ can be satisfied as described above. We assume that σ is negative (i.e., a pressure) all over the interface so that $v' = 0$ can be applied in the entire contact region.

The boundary conditions within the contact zone, i.e., $Y = 0$ and $|X| \le 1$, are

$$|X| \le C: \quad u'(X) = 0, \; v'(X) = 0, \; |\tau(X)| \le \mu|\sigma(X)|, \quad (7a,b,c)$$

$$C < |X| \le 1: \quad v'(X) = 0, \; |\tau(X)| = \mu|\sigma(X)|. \quad (7d,e)$$

Because of the symmetry of $\sigma(X)$ and antisymmetry of $\tau(X)$ and by use of the boundary conditions ($7a,b,d$) we are able to write (1), ($7e$), and ($7c$) as

$$2\int_0^1 \frac{S\tau(S)}{X^2 - S^2}\, dS = -C_1\sigma(X) - C_3\sigma_0, \quad 0 \le X \le C, \quad (8a)$$

$$2X\int_0^1 \frac{\sigma(S)}{X^2 - S^2}\, dS = -C_4\tau(X), \quad 0 \le X \le 1, \quad (8b)$$

$$\tau(X) = \mu\sigma(X), \quad C < X \le 1, \quad (8c)$$

$$|\tau(X)| \le \mu|\sigma(X)|, \quad 0 \le X \le C. \quad (8d)$$

To get equations suitable for numerical solution the two first expressions of equations (8) are transformed into nonsingular integral equations of the Volterra type. This is feasible if we write T for X and apply the operators

$$\int_0^X \frac{dT}{\sqrt{X^2 - T^2}} \quad \text{and} \quad \int_0^X \frac{T\,dT}{\sqrt{X^2 - T^2}}$$

to ($8a$) and ($8b$), respectively. The equations then read

$$\int_X^1 \frac{q(S)}{\sqrt{S^2 - X^2}}\, dS - \gamma_1 \int_0^X \frac{p(S)}{\sqrt{X^2 - S^2}}\, dS = -\gamma_2\tilde{\sigma}_0, \quad 0 \le X \le C, \quad (9a)$$

$$\int_X^1 \frac{Sp(S)}{\sqrt{S^2 - X^2}}\, dS + \gamma_3 \int_0^X \frac{Sq(S)}{\sqrt{X^2 - S^2}}\, dS = 1, \quad 0 \le X \le 1, \quad (9b)$$

$$q(X) = \mu p(X), \quad C < X \le 1, \quad (9c)$$

$$|q(X)| \le \mu|p(X)|, \quad 0 \le X \le C, \quad (9d)$$

where $p(X)$, $q(X)$, and $\tilde{\sigma}_0$ are defined by ($6a,b,d$), and

$$\gamma_1 = \frac{C_1}{\pi}, \quad \gamma_2 = \frac{C_3}{2}, \quad \gamma_3 = \frac{C_4}{\pi}. \quad (10a,b,c)$$

A numerical solution of (9) is given by Spence (1973) for the case of isotropic material and $\sigma_0 \equiv 0$. We use the same numerical method as Spence to find the unknown distributions p, q and the proper value of $\tilde{\sigma}_0$. The interval of integration in ($9a,b$) is divided into subintervals. In each subinterval the functions $p(X)$ and $q(X)$ are approximated by constant values. The equation ($9c$) is applied in the selected region of slip. In order to get the proper value of $\tilde{\sigma}_0$ as a part of the solution we use the equality, in condition ($9d$), on the subinterval nearest to the slip zone. The validity of the solution, i.e., the assumption of only two slip zones in the area of contact, is checked with the inequality of equation ($9d$). The numerical method is described in more detail in the Appendix.

6 Numerical Results

Here we present numerical results for the indentation problem, with limiting friction, stated in Section 5. These results are calculated with the assumption of plane strain. Further, we assume that the orthotropic half-plane consists of fiber-reinforced material. The fibers are uniformly distributed and oriented in the x-direction, i.e., a transversely isotropic medium with identical elastic properties in the y and z-directions. The calculations are made with the following values of material parameters (using standard notation): $E_x/E_y = 5$, $E_y/G_{xy} = 2$, $\nu_{xy} = 0.3$, and $\nu_{zy} \equiv \nu_{yz} \equiv \nu_{yy} \equiv 0.3$.

Three graphs ($\mu = 0.2, 0.4, 0.6$) showing the relation between C and $\tilde{\sigma}_0$, are presented in Fig. 3. Figures 4 and 5 show pressure and shear stress distributions, respectively, for $\mu = 0.4$.

By studying the ratio q/p within the interval $0 \le X \le C$, for different values of $\tilde{\sigma}_0$ and by using the inequality ($9d$), we can estimate an upper limit of $\tilde{\sigma}_0$. As an illustrative example dis-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

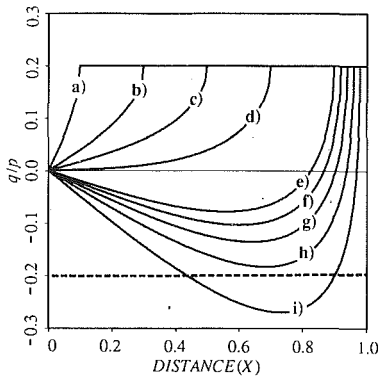


Fig. 6 Distributions of the ratio q/p within the interval $0 \le X \le 1$ ($\mu = 0.2$): a) $\tilde{\sigma}_0 = 0.08$ $C = 0.10$, b) $\tilde{\sigma}_0 = 0.41$ $C = 0.30$, c) $\tilde{\sigma}_0 = 0.60$ $C = 0.50$, d) $\tilde{\sigma}_0 = 0.79$ $C = 0.70$, e) $\tilde{\sigma}_0 = 1.16$ $C = 0.90$, f) $\tilde{\sigma}_0 = 1.23$ $C = 0.92$, g) $\tilde{\sigma}_0 = 1.33$ $C = 0.94$, h) $\tilde{\sigma}_0 = 1.47$ $C = 0.96$, i) $\tilde{\sigma}_0 = 1.72$ $C = 0.98$

tributions of the ratio q/p are calculated for $\mu = 0.20$, and as Fig. 6 shows an upper limit for $\tilde{\sigma}_0$ then lies within the interval 1.47 to 1.72.

7 Concluding Remarks

The size of the two slip zones (i.e., $1 - C$) is decreasing with increasing $\tilde{\sigma}_0$ as shown in Fig. 3. Furthermore, when the coefficient of Coulomb friction (μ) is decreasing a change in $\tilde{\sigma}_0$ will have an increasing influence on the size of the two slip zones. The value of $\tilde{\sigma}_0$ has an upper limit below which there only exist two slip zones, one at each end of the contact zone. Above this upper limit there will be in addition internal slip zones as Fig. 6 indicates. During running conditions for a truck composite leaf spring the value of $\tilde{\sigma}_0$ is within the range where more than two slip zones can be expected in the clamper-leaf spring interface. Therefore, an interesting continuation of this punch problem is to determine similar solutions as considered in this study but with more than two slip zones.

Acknowledgments

I want to thank my supervisor Prof. J. Hult for helpful comments and discussions. This work is partially supported by the Product Development Division at Volvo Truck Corporation at Göteborg (Sweden) and by the Swedish National Board for Technical Development (STU) through the Project No 722-88-02297.

References

Lekhnitsky, S. G., 1937, "Plane Statical Problems of the Theory of Elasticity of an Anisotropic Body," *Prikladnaya matematika i mekhanika,* Vol. 1, No. 1.

Lekhnitsky, S. G., 1981, *Theory of Elasticity of an Anisotropic Elastic Body,* rev. ed., Mir Publisher, Moscow.

Muskhelishvili, N. I., 1949, *Some Basic Problems of the Mathematical Theory of Elasticity,* 3rd ed., Moscow, English translation by J. R. M. Radok, Nordhoff.

Spence, D. A., 1973, "An Eigenvalue Problem for Elastic Contact with Finite Friction," *Proc. Cambridge Philosophical Society,* Vol. 73, pp. 249–268.

Sundelius, B., 1981, "A two-dimensional contact problem with given region of adhesion," *IMA Journal of Applied Mathematics,* Vol. 27, pp. 431–454.

Söhngen, H., 1954, "Zur Theorie der endlichen Hilbert Transformation," *Math. Zeitschrift,* Vol. 60, pp. 31–51.

APPENDIX

A Brief Derivation of the Main Equations by Means of Fourier Transforms

Within a two-dimensional orthotropic homogeneous elastic body the following partial differential equation must be satisfied by Airy's stress function, ϕ (Lekhnitsky, 1981)

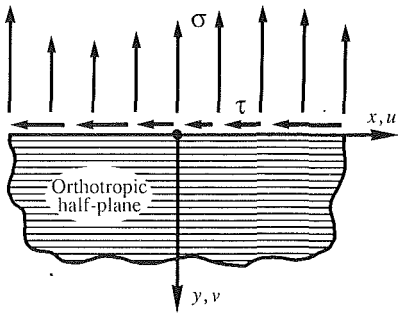


Fig. 7 Orthotropic half-plane loaded along the entire edge by distributions of normal (σ) and shear (τ) stresses

$$\frac{\partial^4 \phi}{\partial y^4} + \frac{c_{66} + 2c_{12}}{c_{11}} \cdot \frac{\partial^4 \phi}{\partial x^2 \partial y^2} + \frac{c_{22}}{c_{11}} \cdot \frac{\partial^4 \phi}{\partial x^4} = 0, \tag{A1}$$

where c_{11}, c_{12}, c_{22}, and c_{66} being the elastic constants in the generalized Hooke's law for an orthotropic medium, where the principal directions of elasticity coincide with the coordinates axes (xyz), i.e.,

$$\epsilon_x = c_{11}\sigma_x + c_{12}\sigma_y, \ \epsilon_y = c_{12}\sigma_x + c_{22}\sigma_y, \ \gamma_{xy} = c_{66}\tau_{xy}. \tag{A2a,b,c}$$

Introduce the following exponential Fourier and inversion transforms of the function $f(x,y)$

$$F(\xi,y) = \int_{-\infty}^{\infty} f(x,y) e^{-i\xi x} dx,$$

$$f(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\xi,y) e^{ix\xi} d\xi, \tag{A3a,b}$$

where i is the imaginary unit. Throughout we shall use a capital letter, say F, to represent the Fourier transform of a function f.

By aid of equation (A3a) we transform (A1) into an ordinary differential equation for the unknown Fourier transform of $\phi(x,y)$, i.e., $\Phi(\xi,y)$

$$\frac{d^4\Phi}{dy^4} - \frac{c_{66} + 2c_{12}}{c_{11}} \cdot \xi^2 \frac{d^2\Phi}{dy^2} + \frac{c_{22}}{c_{11}} \cdot \xi^4 \Phi = 0. \tag{A4}$$

The form of the solution of (A4) depends on the roots of the characteristic equation

$$r^4 - \frac{c_{66} + 2c_{12}}{c_{11}} \cdot \xi^2 r^2 + \frac{c_{22}}{c_{11}} \cdot \xi^4 = 0. \tag{A5}$$

Lekhnitsky (1937) showed that the roots of (A5) cannot be purely imaginary numbers. The following three cases are possible:

I The roots of (A5) are real and unequal

$$r_{1,2} = \pm\alpha|\xi|, \ r_{3,4} = \pm\beta|\xi|, \text{ where } \alpha, \beta > 0. \tag{A6a,b}$$

II The roots are real and equal in pairs

$$r_{1,2,3,4} = \pm\alpha_1|\xi|, \text{ where } \alpha_1 > 0. \tag{A6c}$$

III The roots are complex

$$r_{1,2} = (\alpha_2 \pm i\beta_2)|\xi|, \ r_{3,4} = (-\alpha_2 \pm i\beta_2)|\xi|,$$

$$\text{where } \alpha_2, \beta_2 > 0. \tag{A6d,e}$$

The solution of (A4) for case I, is

$$\Phi(\xi,y) = A \cdot e^{-\alpha|\xi|y} + B \cdot e^{-\beta|\xi|y} + C \cdot e^{\alpha|\xi|y} + D \cdot e^{\beta|\xi|y}, \tag{A7}$$

where the arbitrary constants A, B, C, and D may be determined from the boundary and infinity conditions.

Consider the orthotropic half-plane in Fig. 7, loaded along the entire edge by distributions of normal (σ) and shear (τ) stresses. The boundary conditions at $y = 0$ are then

$$\sigma_y(x,0) = \sigma(x), \ \tau_{xy}(x,0) = \tau(x), \ \forall x, \tag{A8a,b}$$

and the conditions when y tends to infinity are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\sigma_x,\ \sigma_y,\ \tau_{xy}\to 0 \text{ when } y\to\infty. \qquad \text{(A9}a,b,c\text{)}$$

The Fourier transforms of the components of stress σ_x, σ_y, and τ_{xy}, expressed in terms of ϕ, are, respectively,

$$\Sigma_x(\xi,\ y)=\frac{d^2\Phi}{dy^2},\ \Sigma_y(\xi,y)=-\xi^2\Phi, T_{xy}(\xi,y)=-i\xi\frac{d\Phi}{dy}. \qquad \text{(A10}a,b,c\text{)}$$

By aid of (A7) and (A10) it follows from the conditions (A8) and (A9) that

$$A=-\frac{\beta\Sigma(\xi)-i\cdot\text{sgn}(\xi)T(\xi)}{(\beta-\alpha)\xi^2}, B=\frac{\alpha\Sigma(\xi)-i\cdot\text{sgn}(\xi)T(\xi)}{(\beta-\alpha)\xi^2}, \qquad \text{(A11}a,b\text{)}$$

$$C=D\equiv 0, \qquad \text{(A11}c,d\text{)}$$

where $\Sigma(\xi)$ and $T(\xi)$ are the Fourier transforms of $\sigma(x)$ and $\tau(x)$, respectively, and

$$\text{sgn}(\xi)=\begin{cases}+1, & \xi>0,\\ -1, & \xi<0.\end{cases} \qquad \text{(A12)}$$

By substituting the expressions (A11) into (A7), we obtain the final expression for the Fourier transform of the stress function $\phi(x,y)$

$$\Phi(\xi,y)=\frac{1}{(\beta-\alpha)\xi^2}\cdot[(\alpha e^{-\beta|\xi|y}-\beta e^{-\alpha|\xi|y})\Sigma(\xi) + i\cdot\text{sgn}(\xi)\cdot(e^{-\beta|\xi|y}-e^{-\alpha|\xi|y})T(\xi)]. \qquad \text{(A13)}$$

By substituting the Fourier transformed components of strain ϵ_x, ϵ_y, and γ_{xy}, expressed in terms of the displacement u and v (defined in Fig. 7), into the Fourier transformed expressions of (A2) and then substituting Σ_x, Σ_y, T_{xy} from (A10) we obtain

$$U(\xi,y)=-i\cdot\left(\frac{c_{11}}{\xi}\cdot\frac{d^2\Phi}{dy^2}-c_{12}\cdot\xi\Phi\right),$$

$$V(\xi,y)=\frac{c_{11}}{\xi^2}\cdot\frac{d^3\Phi}{dy^3}-(c_{66}+c_{12})\cdot\frac{d\Phi}{dy}, \qquad \text{(A14}a,b\text{)}$$

where U and V are the Fourier transforms of u and v, respectively.

In order to get the expressions of the components of stresses and displacements, we use the Fourier inversion formula (A3b). It is more useful for calculation if we choose to specify the components of displacement gradients $\partial u/\partial x$ and $\partial v/\partial x$, rather than u and v. To obtain the components of edge displacement gradients we let $y=0$ before using the Fourier inversion formula and then use the symmetries in the elastic constants as required by Betti's theorem. We obtain the following result:

The components of stresses within the orthotropic half-plane, i.e., in the region $y\geq 0$ and $|x|\leq\infty$, read

$$\alpha_x(x,y)=\frac{\alpha+\beta}{\pi}\int_{-\infty}^{\infty}\frac{[\alpha\beta y\sigma(s)+(x-s)\tau(s)](x-s)^2}{[(\alpha y)^2+(x-s)^2][(\beta y)^2+(x-s)^2]}ds, \qquad \text{(A15}a\text{)}$$

$$\sigma_x(x,y)=\frac{\alpha+\beta}{\pi}\int_{-\infty}^{\infty}\frac{[\alpha\beta y\sigma(s)+(x-s)\tau(s)](x-s)^2}{[(\alpha y)^2+(x-s)^2][(\beta y)^2+(x-s)^2]}ds, \qquad \text{(A15}a\text{)}$$

$$\sigma_y(x,y)=\frac{\alpha+\beta}{\pi}y^2\int_{-\infty}^{\infty}\frac{[\alpha\beta y\sigma(s)+(x-s)\tau(s)]ds}{[(\alpha y)^2+(x-s)^2][(\beta y)^2+(x-s)^2]}, \qquad \text{(A15}b\text{)}$$

The components of edge ($y=0$) displacement gradients read

$$\frac{du}{dx}(x,0)=(\alpha\beta c_{11}+c_{12})\sigma(x)+\frac{\alpha+\beta}{\pi}c_{11}\int_{-\infty}^{\infty}\frac{\tau(s)}{x-s}ds, \qquad \text{(A16}a\text{)}$$

$$\frac{dv}{dx}(x,0)=\frac{\alpha\beta(\alpha+\beta)}{\pi}c_{11}\int_{-\infty}^{\infty}\frac{\sigma(s)}{x-s}ds - (\alpha\beta c_{11}+c_{12})\tau(x). \qquad \text{(A16}b\text{)}$$

The load σ_0 is not included in the derivation of (A16). However, the only effect of σ_0 will be caused by the additional term $c_{11}\sigma_0$, where $c_{11}\sigma_0$ is the strain ϵ_x ($=\partial u/\partial x$) due to the single load σ_0, in the right-hand side of (A16a).

Formulas for the two other cases of roots can be obtained from (A15) and (A16) by a limiting process on the assumption that (in case II) $\alpha\to\alpha_1$, $\beta\to\alpha_1$, or (in case III) by substituting $\alpha_2+i\beta_2$ and $\alpha_2-i\beta_2$ into α and β.

The General Solution of Coupled Singular Integral Equations of the Second Kind

Consider the equation

$$f(X)+\frac{\lambda}{\pi}\int_{-1}^{1}\frac{f(S)}{X-S}dS=g(X), \quad |X|\leq 1, \qquad \text{(A17)}$$

where $g(X)$, $f(X)$, and λ may be real or complex. The function $g(X)$ is known and it is required to find the function $f(X)$. Here, λ is a parameter whose value depends upon the particular problem. The solution of (A17) is given in the form

$$f(X)=f_1(X)+f_0(X), \quad |X|\leq 0, \qquad \text{(A18)}$$

where $f_0(X)$ is the solution of the homogeneous equation, i.e., equation (A18) with the right-hand side put equal to zero. The functions $f_1(X)$ and $f_0(X)$ read

$$f_1(X)=\frac{g(X)}{1+\lambda^2}-\frac{\lambda}{1+\lambda^2}\cdot\frac{1}{\pi\sqrt{1-X^2}}\left(\frac{1+X}{1-X}\right)^{\gamma} \int_{-1}^{1}\sqrt{1-S^2}\left(\frac{1-S}{1+S}\right)^{\gamma}\frac{g(S)}{X-S}dS, \qquad \text{(A19}a\text{)}$$

$$f_0(X)=-\frac{\lambda}{\sqrt{1+\lambda^2}}\cdot\frac{C}{\pi\sqrt{1-X^2}}\left(\frac{1+X}{1-X}\right)^{\gamma}, \qquad \text{(A19}b\text{)}$$

where γ is a complex constant related to γ by

$$e^{2\pi i\gamma}=\frac{i\lambda-1}{i\lambda+1}, \qquad \text{(A20)}$$

restricted so that its real part $\text{Re}(\gamma)$ lies within the interval $-1/2$ to $+1/2$, and

$$C=\int_{-1}^{1}f(X)dX. \qquad \text{(A21)}$$

Through the definition

$$f(X)=\sigma(X)+i\sqrt{\frac{C_4}{C_1}}\tau(X) \qquad \text{(A22)}$$

it is possible to rewrite the two equations (4) in the form of equation (A17).

Numerical Method

Consider the equations (9). The interval $0\leq S\leq 1$ is divided into N subintervals

$$h_i=\{S:S_{i-1}<S<S_i\},\ i=(1,N), \qquad \text{(A23)}$$

by the points

$$0=S_0<S_1<\ldots<S_N=1. \qquad \text{(A24)}$$

Choose the adhesion boundary C as

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$C = S_M, \text{ where } M < N. \tag{A25}$$

The functions $p(X)$ and $q(X)$ are approximated by piecewise, i.e., in the subintervals, constant functions whose values make up a $2N$ vector $\{p\}$ defined by

$$p_i = p(X_i),\ p_{i+N} = q(X_i),\ i = (1,N), \tag{A26a,b}$$

where X_i are the midpoints of the subintervals,

$$X_i = \frac{1}{2}(S_i + S_{i-1}),\ i = (1,N). \tag{A27}$$

With these assumptions, we rewrite equations (9) as $2N+1$ linear equations of the $2N+1$ unknowns $\{p\}$ and $\tilde{\sigma}_0$

$$\sum_{j=1}^{N}(\gamma_1 A_{ij} p_j - B_{ij} p_{j+N}) = \gamma_2 \tilde{\sigma}_0,\ i = (1,M), \tag{A28a}$$

$$\sum_{j=1}^{N}(C_{ij} p_j + \gamma_3 D_{ij} p_{j+N}) = 1,\ i = (1,N), \tag{A28b}$$

$$p_{i+N} - \mu p_i = 0,\ i = (M+1,N), \tag{A28c}$$

$$p_{M+N} - \mu p_M = 0, \tag{A28d}$$

where

$$A_{ij} = \begin{cases} 0, & j > i, \\ \displaystyle\int_{S_{j-1}}^{X_i} \frac{dS}{\sqrt{X_i^2 - S^2}} = \frac{\pi}{2} - \arcsin\left(\frac{S_{j-1}}{X_i}\right), & j = i, \\ \displaystyle\int_{S_{j-1}}^{S_j} \frac{dS}{\sqrt{X_i^2 - S^2}} = \arcsin\left(\frac{S_j}{X_i}\right) - \arcsin\left(\frac{S_{j-1}}{X_i}\right), & j < i, \end{cases} \tag{A29a}$$

$$B_{ij} = \begin{cases} \displaystyle\int_{S_{j-1}}^{S_j} \frac{dS}{\sqrt{S^2 - X_i^2}} = \ln\left(\frac{S_j + \sqrt{S_i^2 - X_i^2}}{S_{j-1} + \sqrt{S_j^2 - X_i^2}}\right), & j > i, \\ \displaystyle\int_{X_i}^{S_j} \frac{dS}{\sqrt{S^2 - X_i^2}} = \ln\left(\frac{S_j + \sqrt{S_j^2 + X_i^2}}{X_i}\right), & j = i, \\ 0, & j < i, \end{cases} \tag{A29b}$$

$$C_{ij} = \begin{cases} \displaystyle\int_{S_{j-1}}^{S_j} \frac{SdS}{\sqrt{S^2 - X_i^2}} = \sqrt{S_j^2 - X_i^2} - \sqrt{S_{j-1}^2 - X_i^2},\ j > i, \\ \displaystyle\int_{X_i}^{S_j} \frac{SdS}{\sqrt{S^2 - X_i^2}} = \sqrt{S_j^2 - X_i^2}, & j = i, \\ 0, & j < i, \end{cases} \tag{A29c}$$

$$D_{ij} = \begin{cases} 0, & j > i, \\ \displaystyle\int_{S_{j-1}}^{X_i} \frac{SdS}{\sqrt{X_i^2 - S^2}} = \sqrt{X_i^2 - S_{j-1}^2}, & j = i, \\ \displaystyle\int_{S_{j-1}}^{S_j} \frac{SdS}{\sqrt{X_i^2 - S^2}} = \sqrt{X_i^2 - S_{j-1}^2} - \sqrt{X_i^2 - S_j^2}, & j < i. \end{cases} \tag{A29d}$$

The solutions of p and q show (see Figs. 4 and 5) that it is advisable to use different sizes of the subintervals (h_i). A refining division has been used in regions where the slope of the distributions of p and q is relatively high, i.e., in the regions close to $|X| = C$ and $|X| = 1$. Up to 476 subintervals (N) have been used. Equations (A28) are solved by using the numerical routine F04ATF (Mark 13) from the Fortran library NAG (Numerical Algorithms Group). These calculations have been performed on a VAX-station.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Complementarity Problem Formulation of Three-Dimensional Frictional Contact

B. M. Kwak
Department of Mechanical Engineering,
Korea Advanced Institute of
Science and Technology,
Seoul, Korea 135
Mem. ASME

A general three-dimensional frictional contact problem has been formulated in the form of a complementarity problem in an incremental analysis setting. The derivation is straightforward and very natural. It is shown that the complementarity problem for a three-dimensional case is inherently nonlinear, not like the two-dimensional problem where a linear complementarity problem formulation is possible. The two-dimensional case is a special case of the three-dimensional formulation. Approximate linear complementarity problems of the nonlinear complementarity problem by a Newton approach, or by introducing polyhedral law of friction, have been proposed for efficient numerical implementations.

1 Introduction

Contact problems have long been a topic of intensive study since the Hertzian theory developed in 1881, mostly through the analytical approach (Johnson, 1985) before the 1970's and later through computational studies with the advent of high-speed computers (Conry and Seireg, 1971; Chan and Tuba, 1971; Francavilla and Zienkiewicz, 1975; Chand, Haug, and Rim, 1976). The works by analytical methods, including the Hertzian theory, are often limited to regular geometries with elastic material where fundamental singular solutions have been used. The Hertzian theory, however, has had significant applications to the rail-wheel mechanics and rolling element design and analysis.

In solid mechanics, the contact problem is one of the most difficult topics dealing with highly nonlinear phenomena showing both geometric and material, even in the case with linearly elastic material. The geometric nonlinearity comes from the unknown kinematic boundary condition; that is, the unknown contact area, while the material nonlinearity comes from the frictional property.

For the contact without friction general theories are available and very applicable for efficient numerical treatments (Haug and Kwak, 1978; Park and Kwak, 1986). For a large deformation, nonlinear material problem where an incremental formulation is used, the same theories can be naturally applicable (Lee and Kwak, 1984; Joo and Kwak, 1986; Lee and Kwak, 1989). The main theory is that an equivalent optimization exists: the minimization of the potential energy under some kinematic constraints. This can then be formulated and interpreted in several equivalent formulations such as linear complementarity formulation, complementary energy formulation, variational inequality formulation, and so on.

In the friction case, the nonlinearity also involves irreversible and dissipative nature and hence it is loading-path dependent as is well known (Klarbring, 1986b; Kwak and Lee, 1988; Martins and Oden, 1983). Therefore, few theories have been available for this difficult but important frictional contact problem. Even the few theories appearing in the literature still have significant restrictions, or they are able to deal with only a limited scope of problems.

One of the earliest theoretical formulation of frictional problems is a variational inequality (Duvaut and Lions, 1976) and still very prevalent. It is, however, limited to the case where either the normal contact force is assumed known or the contact area is constant. The resulting variational inequality contains a nonsmooth functional corresponding to the frictional work. This formulation is taken further by several authors (Oden and Pires, 1983; Kalker, 1988; Klarbring, 1988) and well explained in Panagiotopoulos (1985). The variational inequality problem, with the normal contact force taken as an unknown variable, is left open.

Considering the mathematical difficulties with the classical Coulomb law of friction, Oden and Pires (1983) presented the so-called nonlocal, nonclassical friction law. This leads to a variational inequality formulation with a smoothing operator mollifying the difficult normal contact pressure. Another formulation of the frictional contact is a complementarity problem formulation (Klarbring, 1986b; Kwak and Lee; 1988) as described in Klarbring (1988), where a subdifferential analysis is used. The friction force and displacement relationship is seen similar to a generalized constitutive law as for elastic-plastic material. For a proper formulation in this direction, they have taken the contact area as constant or the normal displacement on the potential contact region is assumed to be prescribed (Panagiotopoulos, 1985, Klarbring, 1988). In ad-

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, June 6, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

dition, most of the studies described above, except Klarbring (1986a), refer to only two-dimensional cases. Also, they all have dealt with only isotropic friction law.

As for numerical analyses, one approach is to discretize the theoretical formulations such as those described previously. Another approach is to use a direct, heuristic iteration method based on some intuitive systematic trial-and-error method to satisfy the friction condition and geometric compatibility in a discretized formulation (Tsuta and Yamaji, 1973; Herrmann, 1978; Pian and Kubomura, 1981; Rahman et al., 1984; Bathe and Chaudhary, 1985; Mehlhorn et al., 1985; Chandreasekaran et al., 1987; and so on). Contact or interface elements are often used in this approach.

Recently the author (Kwak and Lee, 1988) presented a linear complementarity problem formulation which is directly derived from contact compatibility and friction conditions. It is rather straightforward and simple. Moreover, this new derivation is very natural and generally applicable, as compared with those in the literature. The boundary element method (BEM) is used, which again is very natural because the contact problem is in itself a problem of surface as is the BEM. Here, the formulation is confined only to two-dimensional problems. An extension of the same concept to three-dimensional problems has not been at all obvious. As will be shown soon, however, the two-dimensional formulation is a special case of the three-dimensional formulation proposed here.

Study of the uniqueness of a solution, or finding the correct path of deformation by the incremental calculations, is found to be very important. However, no such study is made in the literature. It is worthwhile to note that in the literature most of the results about uniqueness for a friction problem is for cases with special restrictions, such as with a negligibly small friction coefficient and so on (Demkowicz and Oden, 1982, Klarbring, 1986a). As a matter of fact, as noted already, there should not be a unique solution to a given level of external load. Instead, the state should be uniquely determined for a given path of loading. Therefore, the problem is how to guarantee the correct path numerically. This will restrict the amount of an incremental step for which a unique solution is guaranteed. This topic will be addressed elsewhere with a numerical implementation.

In this paper, a complementarity problem formulation is presented for general contact problems with the Coulomb friction. It is general, because it covers full three-dimensional contacts with an orthotropic friction law. Also, there is no difference whether a multibody contact or a contact against a rigid body is considered. Another special feature of this method is the inclusion of rigid body degree-of-freedoms (Chand, Haug and Rim, 1976; Haug and Kwak, 1978), which is not usually considered in the literature although it is very important in practice. Since an incremental approach is inevitable for the treatment of path dependency of friction effect, usual nonlinearities, such as large deformation and elastoplasticity, can be easily incorporated.

2 Problem Formulation

The proposed formulation is described for a large deformation, three-dimensional contact problem with orthotropic friction law of the Coulomb type. An incremental analysis is inevitable because of loading-path dependency. In later sections, the formulation will be specialized for simpler contact problems, and compared with existing formulations in the literature.

In the incremental approach the configuration, and state up to time t (or a timelike parameter), are known, and those at time $t + \Delta t$ are to be sought. The updated Lagrangian approach is taken for illustration. In this method, all the static and kinematic state variables are referred to the configuration at time t. Following the usual notation in texts, let the coordinates

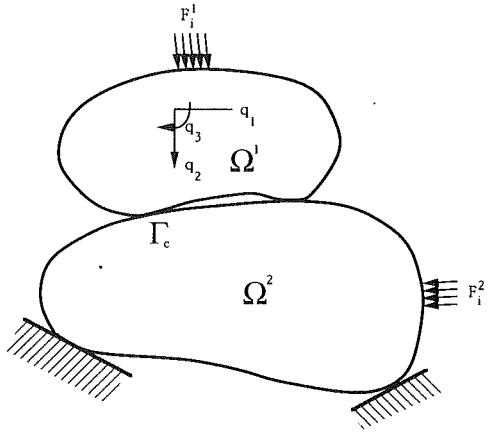


Fig. 1 Two bodies in contact

of a generic point in a body be denoted as a_i and x_i at time t and $t + \Delta t$, respectively. The displacement increment during Δt is then,

$$u_i = x_i - a_i. \tag{1}$$

The energy-conjugate strain and stress measure, referred to the configuration at t, are the Lagrangian strain tensor, ϵ_{ij}, and the second Piola-Kirchhoff stress tensor, ${}^{t+\Delta t}S_{ij}$. They are described by

$$\epsilon_{ij} = e_{ij} + \eta_{ij}, \tag{2}$$

where

$$e_{ij} = (u_{i,j} + u_{j,i})/2 \text{ and } \eta_{ij} = u_{k,i}u_{k,j}/2. \tag{3}$$

The derivative $(\ ),_j$ means differentiation with respect to the coordinate a_j. The second Piola-Kirchhoff stress tensor has the following relationship with the Cauchy stress tensor, ${}^{t+\Delta t}\sigma_{ij}$, which represents the stress at $t + \Delta t$ and is referred to the configuration at $t + \Delta t$. That is,

$${}^{t+\Delta t}\sigma_{ij} = |J| x_{i,r}\, {}^{t+\Delta t}S_{rs} x_{j,s} \tag{4}$$

where $|J|$ is the determinant of the Jacobian of the coordinate transformation. The second Piola-Kirchhoff stress is decomposed as follows

$${}^{t+\Delta t}S_{ij} = {}^{t}\sigma_{ij} + S_{ij}, \tag{5}$$

where S_{ij} denotes the second Piola-Kirchhoff stress increment tensor.

The proposed formulation is described for a two-body contact without loss of generality. Body 2 is assumed to be restrained against any rigid body displacements, while body 1 is allowed for rigid body displacement, denoted by q_i, suitably defined with respect to a point or a set of points, such that the displacement field of body 1 can be described by the sum of q_i and the relative displacement is described with respect to the reference point or set. This situation is completely general and is shown in Fig. 1. The case of a multibody contact requires one to set up similar equations for each additional body.

For the purpose of description, let the region occupied by the bodies by Ω^1 and Ω^2 and their boundaries be Γ^1 and Γ^2, respectively. Each boundary is composed of three disjoint parts: Γ_u where displacement is prescribed, Γ_f where traction conditions are given, and Γ_c, which is the so-called potential contact region taken to be sufficiently large to cover the real contact area after deformation.

The normal contact stress and tangential contact stress are denoted by S_n and S_t. For body k, $k = 1$ or 2:

$${}^{t}S_n^{\,k} = {}^{t}\sigma_{ij}^{\,k}\, {}^{t}n_i^{\,k}\, {}^{t}n_j^{\,k} \tag{6}$$

$${}^{t}S_{ti}^{\,k} = {}^{t}\sigma_{ij}^{\,k}\, {}^{t}n_j^{\,k} - {}^{t}S_n^{\,k}\, {}^{t}n_i^{\,k} \tag{7}$$

where n and t denote the unit normal and tangential vector,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

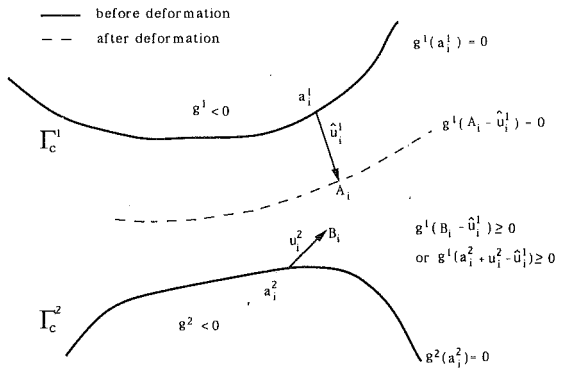


Fig. 2 General impenetration condition

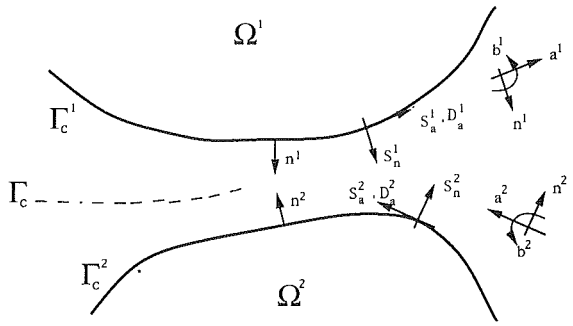


Fig. 3 Principal orthotropic axes and force and relative displacement components

respectively. Let P denote contact pressure taken to be positive by tradition. Then,

$$ {}^tP = -{}^tS_n{}^1 = -{}^tS_n{}^2. \tag{8} $$

The relations, to be satisfied in an incremental step, may now be described as follows.

(2.1) Global Equilibrium. For body 1, where a rigid body motion is allowed, all the external forces and the contact forces should be in equilibrium. For configuration at $t + \Delta t$, from the principle of virtual displacements,

$$ \int_{t+\Delta t\Gamma_f{}^1} {}^{t+\Delta t}F_i H_{ij}{}^{t+\Delta t} d\Gamma + \int_{t+\Delta t\Gamma_c{}^1} {}^{t+\Delta t}S_i A_{ij}{}^{t+\Delta t} d\Gamma = 0, \tag{9} $$

where the coefficient matrices A_{ij} and H_{ij} represent the rigid body displacements of point ${}^{t+\Delta t}\Gamma_c$ and point ${}^{t+\Delta t}\Gamma_f$, respectively, in the ith coordinate direction due to a unit displacement in the jth rigid body degree-of-freedom. The notations F_i and S_i denote traction vectors corresponding to the external and contact forces, respectively. If there is a body force field f_i, an integral over the domain with a kinematic matrix similar to H_{ij} or A_{ij} can be added. For notational simplicity, however, it will be suppressed throughout the paper.

2.2 Internal Equilibrium of Each Body. The local equilibrium equations are expressed in terms of the second Piola-Kirchhoff stress tensor at configuration $t + \Delta t$ as follows (Joo and Kwak, 1986):

$$ ({}^{t+\Delta t}S_{jk} x_{i,k})_{,j} = 0 \quad \text{in } \Omega^k. \tag{10} $$

2.3 Strain-Displacement Relationship. This is already given by equations (2) and (3).

2.4 Stress-Strain Relationship for an Incremental Step.

$$ S_{ij} = C_{ijrs}\epsilon_{rs} \tag{11} $$

where C_{ijrs} is the constitutive coefficient at current time t.

2.5 Boundary Conditions for Each Body. Displacement conditions are

$$ u_i = U_i \quad \text{on } \Gamma_u^k \tag{12} $$

where U_i is the given displacement increment. The traction boundary conditions are

$$ {}^{t+\Delta t}\sigma_{ij}{}^{t+\Delta t}n_j = {}^{t+\Delta t}F_i \quad \text{on } \Gamma_r^k. \tag{13} $$

2.6 Compatibility Condition. No material particle will penetrate the surface of the opposing body. To describe this condition under a very general setting, let the opposing surfaces in the potential contact region at configuration t be denoted by smooth functions,

$$ g^1(a_i^1) = 0 \quad \text{and } g^2(a_i^2) = 0. \tag{14} $$

The interior of each body in the neighborhood of the potential contact region is taken to belong to the negative side of the surfaces as shown in Fig. 2. The impenetration condition says that all the particles a_i^2 should be outside body 1 at configuration $t + \Delta t$. This can be expressed as, for any point a_i^1,

$$ g^1(a_i^2 + u_i^2(a_i^2) - \hat{u}_i^1(a_i^1)) \geq 0 \tag{15} $$

for all a_i^2 satisfying $g^2(a_i^2) = 0$ on $\Gamma_c{}^2$ where

$$ \hat{u}_i^2(a_i^1) = u_i^1(a_i^1) + A_{ij}q_j. \tag{16} $$

It is noted that $g^1(a_i^2 + u_i^2 - \hat{u}_i^1(a_i^1))$ essentially denotes the gap between points a_i^1 and a_i^2 after deformation during Δt.

To describe statement (15) in a simpler term, let

$$ D_n(a_i^1) = \min g^1(a_i^2 + u_i^2(a_i^2) - \hat{u}_i^1(a_i^1)) \tag{17} $$

such that the value $D_n(a_i^1)$ denotes the distance from the particle a_i^1 to the surface $g^2(a_i^2) = 0$ after deformation. The point a_i^1 and the minimizing point a_i^2 can then be called the contacting pair. Then, the impenetration condition can be simply expressed as

$$ D_n(a_i^1) \geq 0 \text{ for all } a_i^1 \text{ such that } g^1(a_i^1) = 0. \tag{18} $$

The compatibility condition can now be described by,

$$ {}^{t+\Delta t}P\, D_n(a_i^1) = 0 \text{ for all } a_i^1 \text{ on } \Gamma_c{}^1. \tag{19} $$

This states that if there is nonzero pressure, then the gap must be zero, and vice versa. It is noted that, instead of finding a_i^2 exactly, an approximate contacting pair obtained from the configuration t is often used. Also noted is that another condition symmetric to equation (15) can be stated with respect to $g^2(a_i^2) = 0$. This one, however, is redundant.

2.7 Friction Conditions. For generality, the following orthotropic friction law of the Coulomb type is considered. For a two-dimensional contact problem there is just one axis and no difference between isotropic and orthotropic law.

Let the principal orthotropic axes on the tangent plane at a point on $\Gamma_c{}^1$ be denoted locally by a and b with the corresponding coefficients of friction, μ_a and μ_b. The tangential traction components and the relative tangential displacement components along the axes are S_a and S_b, and D_a and D_b, respectively. The form of the static friction law is the following:

$$ \left[\left(\frac{{}^{t+\Delta t}S_a}{\mu_a} \right)^2 + \left(\frac{{}^{t+\Delta t}S_b}{\mu_b} \right)^2 \right]^{\frac{1}{2}} \leq {}^{t+\Delta t}P. \tag{20} $$

If strict inequality holds, there is no slip. Otherwise, a relative motion is imminent. In this latter case, it is known (Panagiotopoulos, 1985) that there exists a non-negative λ such that

$$ D_a = -\lambda \frac{{}^{t+\Delta t}S_a}{\mu_a{}^2} \text{ and } D_b = -\lambda \frac{{}^{t+\Delta t}S_b}{\mu_b{}^2} \tag{21} $$

where ${}^{t+\Delta t}P$ is the contact pressure which is non-negative by convention. The relative displacements are defined from the tangential components of the vector $(u^2 - u^1)$ when body 1 is under focus. For description it is convenient to introduce local coordinate systems such as $(a, b, n)^1$ and $(a, b, n)^2$ in Fig. 3.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

It is noted that the axes b^1 and b^2 are in the same direction while the others are opposite each other. With respect to these coordinate systems, all the tangential components are defined. From the action-reaction principle,

$$S_a = S_a^1 = S_a^2,\ S_b = S_b^1 = -S_b^1,\ S_n = S_n^1 = S_n^2$$
$$D_a = D_a^1 = D_a^2,\ D_b = D_a^1 = -D_b^1 \text{ and so on.} \tag{22}$$

Here and in the following, variables without superscripts on the right are referred to body 1 and its local coordinates.

For a two-dimensional case, the Coulomb law says

$$-\mu^{t+\Delta t}P \le {}^{t+\Delta t}S_t \le \mu^{t+\Delta t}P. \tag{23}$$

When strict inequalities hold for both the equations, there is no relative motion: $D_t = 0$. Otherwise, slip is possible: $D_t = -\lambda S_t$ with $\lambda \ge 0$.

It is natural to transform the above statements directly to a so-called complementarity problem form (Cottle et al., 1980). Some complementarity property of the friction law has been recognized and utilized in the literature (Kwak and Lee, 1988; Klarbring, 1988), although limited in scope. The formulation proposed here is very general and presented for orthotropic law of friction. First introduce a non-negative real number ${}^{t+\Delta t}P^* = {}^tP^* + P^*$ and ${}^{t+\Delta t}\theta_f$ such that

$$\frac{{}^tS_a + S_a}{\mu_a} = {}^{t+\Delta t}P^* \cos {}^{t+\Delta t}\theta_f \text{ and } \frac{{}^tS_b + S_b}{\mu_b} = {}^{t+\Delta t}P^* \sin {}^{t+\Delta t}\theta_f. \tag{24}$$

Also introduce D_t and θ_u such that

$$D_a = D_t \cos\theta_u \text{ and } D_b = D_t \sin\theta_u, \tag{25}$$

where D_t denotes the magnitude of slip or the relative displacement between a pair of contacting points. Then the friction condition (20) can be expressed as

$$({}^tP^* + P^*) + T = {}^tP + P \tag{26}$$

with the complementarity rule,

$$T D_t = 0, \tag{27}$$

where $T \ge 0$ is a slack variable, and $D_t \ge 0$. For the two-dimensional case, a similar complementarity condition has already been developed in Kwak and Lee (1988), but it is not directly derivable from the aforementioned three-dimensional statement. The connection between the two will be discussed later. Another relation is derived from equation (21) using equations (24) and (25) as follows:

$$\tan\theta_u = \frac{\mu_a}{\mu_b} \tan {}^{t+\Delta t}\theta_f. \tag{28}$$

This indicates that the direction of tangential contact force can be different from the direction of slip for the orthotropic law.

3 Nonlinear Complementarity System

The governing equations derived may be interpreted as follows. Equations (10), (2), (3), and (11) with the boundary conditions (12) and (13) constitute the usual set of governing equations for deformation mechanics. They essentially provide a relationship between the contact traction vector and displacement vector on the potential contact region Γ_c. Conceptually, this relationship can be expressed as follows:

$$u_c^k = u_c^k({}^{t+\Delta t}S_c^k;\ {}^{t+\Delta t}F^k),\ k = 1,2, \tag{29}$$

where the subscript c refers to the potential contact region Γ_c^k; and ${}^{t+\Delta t}S_c = {}^tS_c + S_c$ denotes a traction vector with both normal and tangential components and can be expressed by ${}^{t+\Delta t}S_a$, ${}^{t+\Delta t}S_b$, ${}^{t+\Delta t}S_n$. An explicit form of this equation in case of a two-dimensional contact is given in (Kwak and Lee, 1988), where the boundary element method has been used. It is noted that an implicit relationship, say, from the principle of minimum energy is also good for the following development.

The global equilibrium equation (9) can be expressed symbolically,

$$E({}^tS_c + S_c;\ {}^tF + F) = 0. \tag{30}$$

Also, if equation (29) is substituted into equations (16) and (17), the gap function can be expressed in the following functional form

$$D_n({}^ta_i^1) = G^1({}^tS_c + S_c,\ g^+,\ g^-;\ {}^ta_t^1,\ {}^ta_i^2), \tag{31}$$

where a_i^1 and a_i^2 are mating contact points predetermined from equation (17) at configuration t, and the rigid body displacement is expressed as a difference of two non-negative numbers, $q = q^+ - q^-$. Then, the condition in equation (19) becomes

$$({}^tP + P)D_n = 0 \tag{32}$$

and

$${}^tP + P \ge 0 \text{ and } D_n \ge 0. \tag{33}$$

Now, since D_a and D_b are functions of u_c^1 and u_c^2, they can also be expressed in terms of S_a, S_b, and S_n using equation (29). Symbolically then,

$$D_a = D_a(S_a,\ S_b,\ S_n)$$
$$D_b = D_b(S_a,\ S_b,\ S_n). \tag{34}$$

To get a consistent formulation, the following manipulation is introduced for equation (30). Since an equality relation can be equivalently stated by two inequalities, equation (30) can be substituted by

$$E({}^tS_c + S_c;\ {}^tF + F) + V^- = 0 \tag{35}$$

and

$$E({}^tS_c + S_c;\ {}^tF + F) - V^+ = 0, \tag{36}$$

where V^+ and V^- are slack variables for the two inequalities. Then, the following complementarity relations can be stated,

$$V^+ q^+ = 0 \text{ and } V^- q^- = 0 \tag{37}$$

where

$$V^+ \ge 0,\ q^+ \ge 0,\ V^- \ge 0,\ q^- \ge 0. \tag{38}$$

This follows, since actually $V^+ = V^- = 0$. It is noted, however, that by adding equations (35) and (36), $V = V^+ - V^-$ is shown to have the meaning of a constrained force, which is zero when q is allowed to be free.

Theoretically it is possible to solve for the free variables, P^*, S_a, S_b, D_a, D_b, θ_f, and θ_u in terms of T, $P(= -S_n)$, and D_t from equations (24)–(26) and (34), when applied for each point on Γ_c. By substituting these solutions into equations (28), (31), (35), and (36), a formal complementarity problem is finally obtained with the complementarity relations (27), (32), and (37). The unknown variables are identified pairwise as,

$$[V^+,\ q^+;\ V^-,\ q^-;\ \{P,\ D_n,;\ T,\ D_t\}]. \tag{39}$$

This is, however, recognized as a nonlinear complementarity system, even with linearly elastic material, because of equations (24) and (25). For a two-dimensional case, a linear complementarity problem has been possible. This shows that the nature of the three-dimensional frictional problem is essentially different from that of a two-dimensional one, although a natural connection exists as will be shown in the following. It is also observed that it is not amenable to use a direct linearization method since the amounts θ_f and θ_u are not necessarily small, because a sudden change in the direction can occur, especially when a slip occurs from a sticking status.

4 Numerical Implementation by Linear Complementarity Problem

4.1 Linearized Geometric Complementarity. The gap function, equation (17), for a mating contact pair can be expanded to obtain the linearized form as follows,

$$D_n = g^1(a_i^2) + (u_k^2 - u_k^1)\dot{g}_k^1 - A_{kj}q_j\dot{g}_k^1 \tag{40}$$

where

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$g_k^{\,1} = \frac{\partial g^1}{\partial a_k^{\,1}} \mid a_k^{\,1} = a_k^{\,2}.$$

It is seen that by dividing by $|\dot{g}_k|$, one obtains a measure of distance. The first term in equation (40) denotes the initial gap at configuration t. The other two terms represent the normal components of the relative and rigid body displacement, respectively. The simple formula that has been derived intuitively and used extensively in the literature (Chand, Haug, and Rim, 1976; Haug and Kwak, 1978; Lee and Kwak, 1986) is a linearized version of this one.

4.2 Approximate Frictional Conditions for Numerical Implementation.

1 Formal Linearization for a Newton-Type Approach. Since the various unknowns corresponding to a small incremental step are small, except possibly the direction changes, θ_f and θ_u, it is natural to use a linearization for these small variables. If the incremental step is considered still too large, some kind of subiterations may be necessary, especially to satisfy the equilibrium. This is the case with elastoplastic and/or large deformation analysis. Some examples for two-dimensional contact problems are found in Joo and Kwak (1986). The question about subiteration will not be pursued here. In any case, it is possible to assume that relation (29) is already linear in S_c.

Notwithstanding that the directional change for an incremental step can be large, let the friction conditions be formally expanded with respect to increments θ_f and θ_u about ${}^0\theta_f$ and ${}^0\theta_u$:

$$\frac{{}^tS_a + S_a}{\mu_a} = ({}^tP^* + P^*)\cos{}^0\theta_f - ({}^tP^* \sin{}^0\theta_f)\theta_f$$

$$\frac{{}^tS_b + S_b}{\mu_b} = ({}^tP^* + P^*)\sin{}^0\theta_f + ({}^tP^* \cos{}^0\theta_f)\theta_f. \tag{41}$$

Expressing D_a and D_b in terms of u_c, and using equation (29), equation (25) becomes

$$D_a(S_a, S_b, S_n) = D_t \cos{}^0\theta_u - ({}^0D_t \sin{}^0\theta_u)\theta_u$$
$$D_b(S_a, S_b, S_n) = D_t \sin{}^0\theta_u - ({}^0D_t \cos{}^0\theta_u)\theta_u, \tag{42}$$

where ${}^0D_t$ is an initial estimate of D_t. If one applies these equations to as many points as there are on the potential contact region for S_a and S_b, the number of equations are enough for solving S_a, S_b, and θ_f, and θ_u for all these points in terms of P, D_t, and T, since $P^* = {}^tP + P - T - {}^tP^*$ from equation (26).

If ${}^0\theta_f$ and ${}^0\theta_u$ are considered as the directions of total tangential forces and slip at an arbitrary configuration ${}^0t$, and if θ_f and θ_u can be guaranteed small, the above system is just as good as the other linearized systems, and the four sets of variables can be eliminated from equations (35), (36), and (40).

As already noted, however, θ_f and θ_u are not necessarily small as assumed in the last section, because a sudden direction change into the opposite direction is possible. In this case, the above system can be considered as a Newton step, for finding ${}^{t+\Delta t}\theta_f^{(i+1)} = {}^0\theta_f^{(i)} + \theta_f$, where the index i denotes subiteration. Once P^* and D_t are numerically given, this iteration can be performed numerically. Otherwise, the whole complementarity system may be iterated, until a convergence is obtained.

2 Approximation by the Polyhedral Law of Friction. The second and better approach is to replace the original elliptic cone of acceptable traction forces defined in equation (20), by an approximate polyhedral cone. One such replacement appears in Klarbring (1988) in an example for the isotropic law of friction.

To explain in general terms, denote the principal axes by a, b, and normal n. Let C_E and $C_P(e)$ be the original elliptic cone and the polyhedral cone, respectively, as follows:

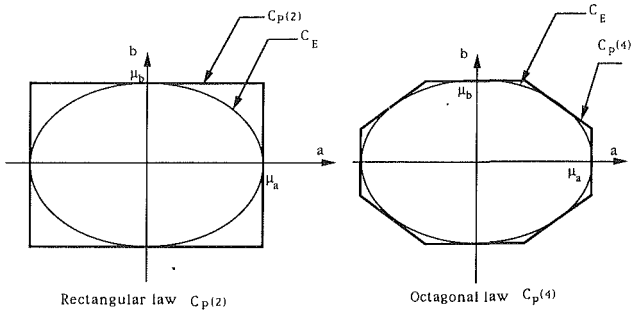


Fig. 4 Polyhedral law of friction as compared to elliptical law

$$C_E = \left\{ (S_a, S_b, P) : \left[\left(\frac{{}^{t+\Delta t}S_a}{\mu_a} \right)^2 + \left(\frac{{}^{t+\Delta t}S_b}{\mu_b} \right)^2 \right]^{\frac{1}{2}} \le {}^{t+\Delta t}P \right\} \tag{43}$$

and

$$C_P(e) = \left\{ (S_a, S_b, P) : \left| \frac{{}^{t+\Delta t}S_a}{\mu_a} \cos\alpha_i + \frac{{}^{t+\Delta t}S_b}{\mu_b} \sin\alpha_i \right| \le {}^{t+\Delta t}P,\ i = 0, \ldots .e \right\} \tag{44}$$

where α_i is a directional angle in the (a, b) plane and e denotes the number of pairs of opposing faces of the polyhedral cone. It is noted that the term on the left-hand side of inequality sign in equation (44) denotes the component of tangential force along the direction α_i. The cross-sectional shapes of the cones, C_E, $C_P(2)$, and $C_P(4)$ for ${}^{t+\Delta t}P = 1$ are shown in Fig. 4. The cone $C_P(1)$ with $\alpha_1 = 0$ corresponds to a one-dimensional case. The cone $C_P(2)$ may be termed as a rectangle law of friction due to its shape. This may be interpreted as the result of direct application of the one-dimensional law along each principal axis as follows:

$$|{}^{t+\Delta t}S_a| \le \mu_a {}^{t+\Delta t}P \text{ and } |{}^{t+\Delta t}S_b| \le \mu_b {}^{t+\Delta t}P. \tag{45}$$

The general case with arbitrary e will be called here as polyhedral law of friction.

To arrive at a complementarity rule for the polyhedral law, introduce slack variables T_i^- and T_i^+ which are non-negative such that

$$\frac{{}^{t+\Delta t}S_a}{\mu_a}\cos\alpha_i + \frac{{}^{t+\Delta t}S_b}{\mu_a}\sin\alpha_i + \frac{T_i^-}{\mu^*} = {}^{t+\Delta t}P$$
$$\frac{{}^{t+\Delta t}S_a}{\mu_a}\cos\alpha_i + \frac{{}^{t+\Delta t}S_b}{\mu_a}\sin\alpha_i - \frac{T_i^+}{\mu^*} = -{}^{t+\Delta t}P, i = 1, \ldots, e \tag{46}$$

where $\mu^* = \sqrt{(\mu_a^2 + \mu_b^2)/2}$ has been introduced for future convenience.

Define the relative displacement D_{ti}, which is split into positive and negative parts D_{ti}^+ and D_{ti}^-, for the corresponding directions β_i, as follows:

$$D_a \cos\beta_i + D_b \sin\beta_i = D_{ti}^+ - D_{ti}^-,\ i = 1, \ldots .e \tag{47}$$

where

$$D_{ti}^+ \ge 0,\ D_{ti}^- \ge 0, \tag{48}$$

and

$$\tan\beta_i = \frac{\mu_a}{\mu_b}\tan\alpha_i. \tag{49}$$

This last relation comes from equation (28). If the component of a tangential force vector $({}^{t+\Delta t}S_a,\ {}^{t+\Delta t}S_b)$ computed along

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the direction α_i falls on the boundary of the polyhedral cone, a slip in the opposite to the force may occur along β_i. That is, when $D_{ti}^- > 0$, T_i^- must be zero, and when $T_i^- > 0$, there can not be a slip in the negative direction. This condition can be stated as

$$T_i^- {}_{ti}D^- = 0,\ i = 1, \ldots, e.$$

Similarly,

$$T_i^+ D_{ti}^+ = 0,\ i = 1, \ldots, e, \tag{50}$$

where

$$T_i^+ \geq 0,\ D_{ti}^+ \geq 0,\ T_i^- \geq 0,\ D_{ti}^- \geq 0.$$

It is reminded that ${}^{t+\Delta t}S_a = {}^tS_a + S_a$, ${}^{t+\Delta t}S_b = {}^tS_b + S_b$ and ${}^{t+\Delta t}P = {}^tP + P$. The system of equations may now be summarized as follows: equations (46), for $i = 1$ and 2, equation (47) for $i = 1$ and 2, equations (34), (35), (36), and (40), and the complementary relations, equations (32), (37), and (50), for $i = 1$ and 2. Among the unknowns, S_a, S_b, D_n, D_a, D_b, P, V^+, V^-, q^+, q^-, D_{ti}^+, D_{ti}^-, T_i^+, and T_i^-, the unknowns S_a, S_b, D_a, D_b can be readily eliminated from equations (34) and (46). If $e > 2$, it is better to impose the condition on frictional force pairwise, such that equations (46), (47), and (50) are imposed with different α_i's. The resulting set of relations with non-negativity conditions constitute a linear complementarity problem, with the primary unknowns as follows:

$$[V^+, q^+; V^-, q^-; \{P, D_n; (T_i^+, D_{ti}^+; T_i^-, D_{ti}^-,\ i = 1, \ldots, e)\}]. \tag{51}$$

It is noted that when $C_P(1)$ is used corresponding to a two-dimensional case, the formulation described in Kwak and Lee (1988) is obtained as a special case of the three-dimensional problem. When $C_P(e)$ is used, the number of complementary variable pairs is $2 + (2e + 1)$ (number of checkpoints on Γ_c), as compared to that of the original variable pairs, $2 + 2$ (number of checkpoints), from expression (39). For numerical implementation it is convenient to deal with equations (46) and (47) for two angles at a time, with 90 deg apart, as noted earlier. This corresponds to imposing a rectangular law of friction one after another.

It is observed that the polyhedral law of friction is simple to use and leads immediately to a linear complementarity problem for the frictional contact problem. If, however, one still wants to solve the problem with the elliptic law, then from the solution with the polyhedral law, an approximate direction for tangential traction vector can be found by calculating the angle from the solution, ${}^{t+\Delta t}S_a {}^{t+\Delta t}S_b$. These solutions must be good estimates to the real directions, and hence can be used as the initial estimates ${}^0\theta_f$ and ${}^0\theta_u$ in the formal linearization approach where good estimates are essential. This approach may be termed as a mixed method, indicating that good initial estimates are first obtained from other methods, and they are used for the original linearized problem.

4.3 Updating Configurations and State Variables. After an incremental step for Δt is processed, the configuration at $t + \Delta t$ is taken as a new reference configuration for the next increment. Correspondingly, the coordinates and the state variables must all be updated, such that $a_i^{\text{new}} = a_i + u_i$, $q_j^{\text{new}} = q_j^{\text{old}} + q_j$, and so on. With the configuration change, the orthotropy direction for friction must also be updated.

5 Reduction to Simpler Cases

5.1 Reduction to Two-Dimensional Frictional Contact. For a two-dimensional case, the Coulomb law says that $|S_t| \leq \mu P$. This can be equivalently stated as,

$$S_t + T^- = \mu P$$
$$-S_t + T^+ = \mu P \tag{52}$$

where slack variables are introduced as before. The relative slip D_t is expressed accordingly as a difference of two non-negative values:

$$D_t = D_t^+ - D_t^- \tag{53}$$

where

$$D_t = D_t(S_t, P; F). \tag{54}$$

Then the following complementarities hold

$$D_t^+ T^+ = 0 \text{ and } D_t^- T^- = 0. \tag{55}$$

By eliminating $\dot{S}_t$ from equations (52),

$$T^+ + T^- = 2\mu P. \tag{56}$$

Also, since $S_t = (T^+ - T^-)/2$, S_t can be eliminated from all other expressions such as for D_n and D_t. Therefore, equations (35), (36), (40), (54), and (56) constitute the complementarity problem sought for with complementarity relations (32), (37), and (55). The resulting system is linear and found to be the same as in Kwak and Lee (1988).

5.2 Reduction to Frictionless Contact. If there is no friction, then the friction condition described in Section 2, is not necessary, and all tangential contact forces are zero. The basic equations are simplified. Equation (29) now becomes,

$$u_c^k = u_c^k({}^{t+\Delta t}S_n^k,\ {}^{t+\Delta t}F^k),\ k = 1,2. \tag{57}$$

From equations (35) and (36),

$$E({}^tS_n + S_n;\ {}^tF + F) + V^+ = 0$$
$$-E({}^tS_n + S_n;\ {}^tF + F) + V^- = 0. \tag{58}$$

Using equation (57), equation (31) becomes

$$D_n({}^ta_i^1) = G^1({}^tS_n + S_n, q^+, q^-;\ {}^ta_i^1, {}^ta_i^2). \tag{59}$$

The complementarity conditions are, with $P = -S_n$,

$$({}^tP + P)D_n = 0 \tag{60}$$

and

$$V^+ q^+ = 0 \text{ and } V^- q^- = 0, \tag{61}$$

with all variables non-negative.

The system, equations (58) to (61), constitutes a complementarity problem. For the particular case of a linear elastic, small-displacement frictionless contact, it is a linear complementarity problem and is equivalent to the quadratic programming problem derived from the principle of minimum potential energy under kinematic constraints. The linear complementarity problem corresponds to necessary conditions of this quadratic programming problem.

6 Summary and Conclusions

It has been shown that a three-dimensional frictional contact problem can be formulated as a complementarity problem, which is inherently nonlinear even for the case of linearly elastic materials. This fact contrasts itself with the linear complementarity problem formulation possible for two-dimensional frictional problem.

The formulation is very general without any specific restrictions such as the assumptions on normal contact force or on the contact area usually made in the literature. It is applicable to nonlinear material and nonlinear geometric problems with orthotropic law of friction. A linearization suitable for a numerical implementation has been possible by introduction of a polyhedral law of friction approximating the original elliptic law. The complementarity problem previously introduced by the author for two-dimensional contact has been shown to be a special case of this linearized three-dimensional formulation.

7 Acknowledgment

This research is partially supported by the Korea Science and Engineering Foundation.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

Bathe, K.-J., and Chaudhary, A., 1985, "A Solution Method for Planar and Axisymmetric Contact Problems," *Int. J. Num. Meth. Engr.*, Vol. 21, pp. 65–88.

Chan, S. K., and Tuba, I. S., 1971, "A Finite Element Method for Contact Problems of Solid Bodies—Part I: Theory and Validation," *Int. J. Mech. Sci.*, Vol. 13, pp. 615–625.

Chand, R., Haug, E. J., and Rim, K., 1976, "Analysis of Unbonded Contact Problems by Means of Quadratic Programming," *J. Optim. Theory and Appl.*, Vol. 20, No. 2, pp. 171–189.

Chandrasekaran, N., Haisler, W. E., and Goforth, R. E., 1987, "A Finite Element Solution Method for Contact Problems with Friction," *Int. J. Num. Meth. Engr.*, Vol. 24, pp. 477–495.

Conry, T. F., and Seireg, A., 1971, "A Mathematical Programming Method for Design of Elastic Bodies in Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, pp. 387–392.

Cottle, R. W., Giannessi, F., and Lions, J-L., 1980, *Variational Inequalities and Complementarity Problems*, John Wiley and Sons, New York, Chapters 7 and 15.

Demkowicz, L., and Oden, J. T., 1982, "On Some Existence and Uniqueness Results in Contact Problems with Nonlocal Friction," *Nonlinear Analysis, Theory, Meth. & Appl.*, Vol. 6, No. 10, pp. 1075–1093.

Duvaut, G., and Lions, J. L., 1976, *Inequalities in Mechanics and Physics*, (English translation), Springer-Verlag, New York.

Francavilla, A., and Zienkiewicz, O. C., 1975, "A Note on Numerical Computation of Elastic Contact Problems," *Int. J. Num. Meth. Engr.*, Vol. 9, pp. 913–924.

Haug, E. J., and Kwak, B. M., 1978, "Contact Stress Minimization by Contour Design," *Int. J. Num. Meth. Engr.*, Vol. 12, pp. 917–930.

Herrmann, L. R., 1978, "Finite Element Analysis of Contact Problems," *ASCE J. Engr. Mech. Div.*, Vol. EM5, pp. 1043–1057.

Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press.

Joo, J. W., and Kwak, B. M., 1986, "Analysis and Application of Elastoplastic Contact Problems Considering Large Deformations," *Computers and Structures*, Vol. 24, No. 6, pp. 953–961.

Kalker, J. J., 1988, "Contact Mechanical Algorithms," *Comm. Appl. Num. Meth.*, Vol. 4, pp. 25–32.

Klarbring, A., 1986a, "A Mathematical Programming Approach to Three-dimensional Contact Problems with Friction," *Computer Meth. Appl. Mech. Engr.*, Vol. 58, pp. 175–200.

Klarbring, A., 1986b, "General Contact Boundary Conditions and the Analysis of Frictional Systems," *Int. J. Solids Structures*, Vol. 22, No. 12, pp. 1377–1398.

Klarbring, A., 1988, "On Discrete and Discretized Non-linear Elastic Structures in Unilateral Contact (Stability, Uniqueness and Variational Principles)," *Int. J. Solids Structures*, Vol. 24, No. 5, pp. 459–479.

Kwak, B. M., and Haug, E. J., 1976, "Parametric Optimal Design," *J. Optimization Theory and Appl.*, Vol. 20, No. 1, pp. 13–35.

Kwak, B. M., and Lee, S. S., 1988, "A Complementarity Problem Formulation for Two-dimensional Frictional Contact Problem," *Computers and Structures*, Vol. 28, No. 4, pp. 469–480.

Lee, B. C., and Kwak, B. M., 1984, "A Computational Method for Elastoplastic Contact Problems," *Computers and Structures*, Vol. 18, No. 5, pp. 757–765.

Lee, G. B., and Kwak, B. M., 1989, "Formulation and Implementation of Beam Contact Problems Under Large Displacement by a Mathematical Programming," *Computers and Structures*, Vol. 31, No. 3, pp. 365–376.

Martins, J. A. C., and Oden, J. T., 1983, "A Numerical Analysis of a Class of Problems in Elastodynamics with Friction," *Computer Meth. Appl. Mech. Engr.*, Vol. 40, pp. 327–360.

Mehlhorn, G., Kollegger, J., Keuser, M., and Komar, W., 1985, "Nonlinear Contact Problems—A Finite Element Approach Implemented in ADINA," *Computers and Structures*, Vol. 21, No. 1-2, pp. 69–80.

Oden, J. T., and Pires, E. B., 1983, "Nonlocal and Nonlinear Friction Laws and Variational Principles for Contact Problems in Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 67–76.

Oden, J. T., and Pires, E. B., 1984, "Algorithms and Numerical Results for Finite Element Approximations of Contact Problems with Non-classical Friction Laws," *Computers and Structures*, Vol. 19, No. 1-2, pp. 137–147.

Panagiotopoulos, P. D., 1985, *Inequality Problems in Mechanics and Applications*, Birkhauser.

Park, S. G., and Kwak, B. M., 1986, "A Semi-anaytical Finite Element Method for Three-dimensional Contact Problems with Axisymmetric Geometry," *Proc. Insst. Mech. Engrs.*, Vol. 200, No. C6, pp. 399–405.

Pian, T. H. H., and Kubomura, K., 1981, "Formulation of Contact Problems by Assumed Stress, Hybrid Elements," in *Nonlinear Finite Element Analysis in Structural Mechanics*, Springer-Verlag, New York, pp. 49–59.

Rahman, M. U., Rowlands, R. E., and Cook, R. D., 1984, "An Iterative Procedure for Finite-Element Stress Analysis of Frictional Contact Problems," *Computers and Structures*, Vol. 18, No. 6, pp. 947–954.

Tsuta, T., and Yamaji, S., 1973, "Finite Element Analysis of Contact Problem," *Theory and Practice in Finite Element Structural Analysis*, Proc. Tokyo Seminar on Finite Element Analysis, pp. 177–194.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Thermal Elastica Theory of Piston-Ring and Cylinder-Bore Contact[1]

D. C. Sun
Department of Mechanical and Industrial Engineering, The Thomas J. Watson School of Engineering, Applied Science, and Technology, State University of New York at Binghamton, Binghamton, NY 13902-6000
Mem. ASME

The contact between piston-ring and cylinder-bore is generally nonuniform, due to mechanical and thermal distortion of the bore and manufacturing imperfection in the shape of the free ring. This nonuniform contact problem is treated in the present work that features arbitrarily specified bore shape and free ring shape, and also includes the thermal deformation of the ring and the effect of gas pressure. The problem is solved with the elastica theory, incorporating thermal elasticity, and a complementarity formulation. Solutions are computed using dimensions, material values, and operating parameters representative of those found in automotive engines.

1 Introduction

A piston ring, when inserted inside a cylinder, must fit the bore closely to perform its function as a seal—preventing the combustion gas from leaking into the crankcase, and the crankcase oil from leaking into the combustion chamber. Besides this functional requirement, the satisfactory operation of the piston ring also demands the contact force between the ring and the bore to be as small as practicable. This is because the magnitude of the contact force is directly related to the friction at the contact, which contributes to a significant portion of the overall energy loss in the engine, and the possibility of lubricating the contact, which is crucial to the engine's service life. In the current practice, the shape of the free ring is designed in such a way that the ring conforms exactly to a circular bore when subject to a uniformly distributed normal contact force around its periphery; and the analyses of friction and lubrication are all based on such an ideal contact state.

But in reality, the contact is not uniform, because the operating condition of the ring is different than its design condition. First, the bolting of the cylinder head to the engine block causes the bore to become lobed. Secondly, the arrangement of cooling jackets around the cylinder generally results in an uneven temperature distribution, and hence a thermal distortion of the bore. Thirdly, the shape of the free ring is seldom found exactly as designed. In addition to the deviations from design, other factors also obscure the contact picture. One is the location of the ring gap, which is no longer immaterial when bore distortion is present. Another is the effect of gas pressure. When the gas pressure acts only on the ring back, it helps to make the ring conform to the bore. But, where the ring separates from the bore, the gas pressure acts on both the ring back and the ring face; then its helpful role is no longer so obvious. From the brief discussion above, it may be clear that a detailed treatment of this nonuniform contact problem is needed to better understand how the piston ring actually works.

There are several published analytical studies of ring-bore contact. A solution for a circular ring conforming to a circular bore can be found in Timoshenko and Lessells (1925), where a formula is given for the *variable* wall thickness of the ring (of rectangular cross-section) needed to produce a uniformly distributed contact force. Another solution was given by Swift (1947) for a ring of *constant* cross-section. The free ring shape found in this often-quoted work has long been the basis of ring design. Having in mind that a used bore could be worn, Truesdell (1961) obtained an exact solution to the uniform contact problem between a ring of constant cross-section and an *elliptical* bore. Recently, Mierbach et al. (1983) included the thermal deformation of the ring in a set of formulae for calculating the free ring shape that would result in uniform contact with a *circular* bore. The contact between a circular ring of constant cross-section and a circular bore is necessarily nonuniform (except for the trivial case of equal radii). An analytical solution to this problem can be found in Feodosyev (1977). While the solution is limited to the case of vanishing end clearance, it reveals many salient properties of the contact problem and can serve as a useful check for elaborate numerical solutions. The problem has also been solved numerically including the effect of shear deformation (Dragoni and Strozzi, 1986). To the author's knowledge, however, there is no published solution of the nonuniform contact problem considering arbitrarily specified bore shape and free ring shape.

[1]The majority of this work was done while the author was affiliated with General Motors Research Laboratories, Warren, Mich.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 2, 1989; final revision, Apr. 2, 1990.

Paper No. 91-APM-8.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

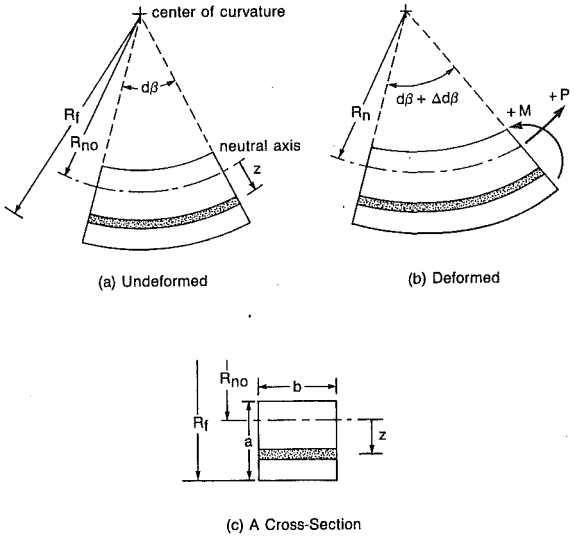


Fig. 1 Notations for a ring element

In this work the general nonuniform ring-bore contact problem is treated. The analysis takes into account arbitrary bore shape, arbitrary free ring shape, thermal deformation of the ring, and the effect of gas pressure. The numerical nature of the analysis also allows a variable ring cross-section to be included, though in the computation only a constant rectangular cross-section is used. The effects of the various features considered are demonstrated by the computed results of sample cases in which the dimensions, material values, and operating parameters are representative of those found in automotive engines.

2 Formulation and Analysis

The contact problem may be divided into two subproblems: one is to find the deflected ring shape from its free shape due to a given set of loads and thermal condition. The other one is to determine these loads so that the deflected ring fits into a cylinder bore of a specified shape. The first problem can be handled with thermoelasticity (Boley and Weiner, 1960) and elastica theory (Frisch-Fay, 1962). The second problem involves the determination of the contact pattern as well as loads, and can be treated as a complementarity problem (Murty, 1976). The problem is simplified by the following assumptions:

(1) The cylinder bore has a given shape, but is rigid.

(2) The bore temperature distribution and the piston temperature are given. The temperature field in the ring has been solved under these conditions, and hence is known.

(3) Ring deflection takes place in a plane orthogonal to the bore axis. As related to this restriction, we consider only in-plane loads and a ring cross-section that is symmetric with respect to the plane.

(4) Loads are normal to the ring, i.e., there is no friction force acting in the circumferential direction of the ring.

(5) While the ring deflection due to a nonuniform temperature field need be considered, the thermal distortion of the ring cross-section may be neglected.

2.1 Deformation of a Ring Element. Consider an element of the free ring, Fig. 1(*a*). We first define a "neutral" axis in terms of its radius of curvature, R_{no}, as follows:

$$\int_A \frac{z dA}{R_{no}+z} = 0 \tag{1}$$

where z is the coordinate measured from the neutral axis, and the integration is over the cross-section, whose area is denoted as A. Note that R_{no} varies along the ring in general, and is therefore a function of the angular coordinate of the free ring, θ_o. From equation (1) we establish the following relations:

$$\int_A \frac{dA}{R_{no}+z} = \frac{A}{R_{no}} \tag{2}$$

and

$$\int_A \frac{z^2 dA}{R_{no}+z} = \int_A z dA \equiv \gamma \int_A dA = \gamma A. \tag{3}$$

For a rectangular cross-section $a \times b$, Fig. 1(*c*), the above relations yield:

$$R_{no} = \frac{a}{\ln\left(\frac{R_f}{R_f - a}\right)} \tag{1'}$$

and

$$\gamma = R_f - R_{no} - \frac{a}{2} \tag{3'}$$

where R_f is the radius of curvature of the ring face.

The length of the marked fiber in Fig. 1(*a*) is $(R_{no} + z)d\beta$, where $d\beta$ is the central angle subtended by the ring element. After deformation, its length becomes $(R_n + z)\ (d\beta + \Delta d\beta)$. See Fig. 1(*b*). Hence, the tensile strain in this fiber is:

$$\epsilon = \frac{1}{(R_{no}+z)d\beta}\left[(R_n+z)(d\beta+\Delta d\beta) - (R_{no}+z)d\beta\right]$$

$$= \frac{1}{R_{no}+z}\left[R_n\left(1+\frac{\Delta d\beta}{d\beta}\right) - R_{no} + z\frac{\Delta d\beta}{d\beta}\right]. \tag{4}$$

But the tensile strain is caused by both a tensile stress σ and a temperature rise T, i.e.,

$$\epsilon = \frac{\sigma}{E} + \alpha T \tag{5}$$

where E is the Young's modulus and α the linear coefficient of thermal expansion. Integrating the tensile stress over the cross-section, one finds the tensile force and bending moment as follows:

$$P = \int_A \sigma dA = \frac{EA}{R_{no}}\left[R_n\left(1+\frac{\Delta d\beta}{d\beta}\right) - R_{no}\right] - E\alpha \int_A T dA \tag{6}$$

$$M = \int_A \sigma z dA = E\gamma A \frac{\Delta d\beta}{d\beta} - E\alpha \int_A T z dA. \tag{7}$$

From equation (7) we obtain:

$$\frac{\Delta d\beta}{d\beta} = \frac{M}{E\gamma A} + \frac{\alpha}{\gamma A}\int_A T z dA. \tag{8}$$

Substituting equation (8) into equation (6), we find:

$$\frac{1}{R_n} = \left(\frac{1}{R_{no}}\right)\frac{1 + \frac{M}{E\gamma A} + \frac{\alpha}{\gamma A}\int_A T z dA}{1 + \frac{P}{EA} + \frac{\alpha}{A}\int_A T dA}. \tag{9}$$

The length of the ring element is defined to be the length of the fiber at the neutral axis. Hence, for the free ring element,

$$ds_o = R_{no} d\beta \tag{10}$$

and for the deformed ring element,

$$ds = R_n(d\beta + \Delta d\beta) = ds_o\left(1 + \frac{P}{EA} + \frac{\alpha}{A}\int_A T dA\right). \tag{11}$$

2.2 Deflection of the Ring. We describe the ring shape by that of its neutral axis. The curves in Fig. 2 represent a

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

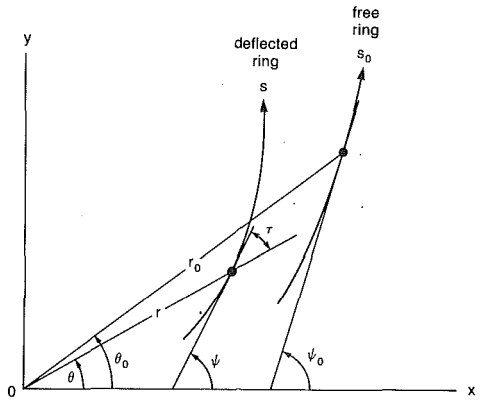


Fig. 2 Notations for ring shape

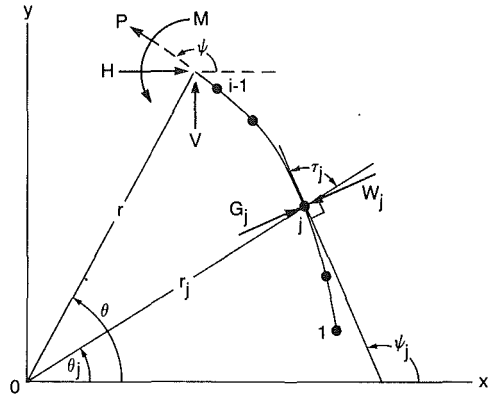


Fig. 3 Moment and forces at a ring cross-section

segment of the deflected ring and the corresponding segment of the free ring. The polar coordinates (r, θ) of a point on the deflected curve are functions of the length coordinate, s. Hence.

$$\frac{dr}{ds} = \cos\tau \tag{12}$$

$$\frac{d\theta}{ds} = \frac{\sin\tau}{r} \tag{13}$$

where τ is the difference between the tangent angle and the angular coordinate,

$$\tau = \psi - \theta. \tag{14}$$

Since $d\psi/ds$ is by definition the curvature, i.e., R_n^{-1}, equation (9) may be written as:

$$\frac{d\psi}{ds} = \left(\frac{d\psi_o}{ds_o}\right) \frac{1 + \dfrac{M}{E\gamma A} + \dfrac{\alpha}{\gamma A}\displaystyle\int_A Tz dA}{1 + \dfrac{P}{EA} + \dfrac{\alpha}{A}\displaystyle\int_A T dA}. \tag{9'}$$

Equations (12), (13), and (9′) govern the deflected ring shape. This system of equations may be integrated from one end of the ring to obtain the deflected shape. To do so, the coordinates and slope of this end are needed. This set of initial values, denoted as c_r, c_θ, c_ψ, are the additional unknowns in the problem.

2.3 Moment and Forces at a Ring Cross-Section. The ring is subject to gas load and contact load. The gas load is assumed to be uniformly distributed over the ring back and also over that part of the ring face where there is gap between the ring and the bore (separation regions). The contact load acts on the part of the ring face where the ring contacts the bore (contact regions). In the interior of the contact regions, the contact load is of distributed nature. However, at the edges of such regions concentrated forces invariably appear, as explained in Feodosyev (1977, pp. 160–166). To avoid the complexity of dealing with both distributed and concentrated loads, and because the problem is to be discretized for numerical treatment any way, we model the loads by concentrated forces acting at node points. Therefore, we discretize the ring into N (an odd integer) nodes, with node 1 and node N at the two ends. The nodes fall in two categories: contact nodes (C) and separation nodes (S). Each node is acted on by an outward pointing gas load G_i and an inward pointing contact load W_i. See Fig. 3. The W_i's at the contact nodes are unknowns in the problem, whereas those at the separation nodes are set to be equal to the corresponding G_i's in magnitudes.

Consider the equilibrium of a ring segment containing nodes 1 to $i-1$. The cut section in Fig. 3 is acted on by a moment M and forces H, V. Let us first determine the forces. The x and y-components of the loads at node j are, respectively:

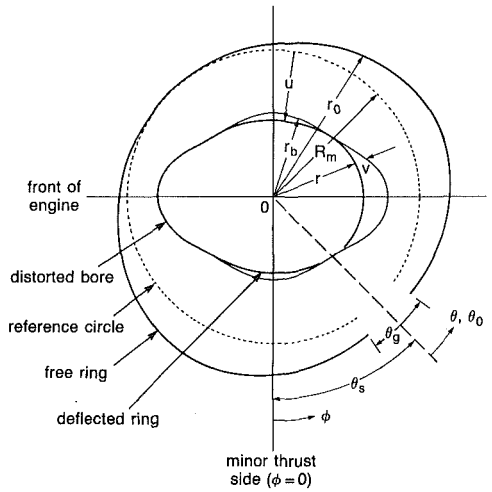


Fig. 4 Notations for contact

$$(W_j - G_j)\cos\left(\psi_j - \frac{\pi}{2}\right)$$

and

$$(W_j - G_j)\sin\left(\psi_j - \frac{\pi}{2}\right).$$

Adding up the contributions of nodes 1 to $i-1$, we obtain:

$$\left.\begin{aligned} H(\theta) &= \sum_{j=1}^{i-1} (W_j - G_j)\sin\psi_j \\ V(\theta) &= -\sum_{j=1}^{i-1} (W_j - G_j)\cos\psi_j \end{aligned}\right\} \tag{15}$$

$$\theta_{i-1} \le \theta < \theta_i,\ i = 2, \ldots, N.$$

Combining the components of H and V in the tangential direction, we obtain the tensile force P at the section:

$$P(\theta) = H\cos\psi + V\sin\psi, \quad \theta_{i-1} \le \theta < \theta_i,\ i = 2, \ldots, N. \tag{16}$$

Let us now determine the moment. By decomposing the loads at node j into two components, one in the direction of the position vector and one perpendicular to it, we find the moment at the cut section due to these components to be:

$$(W_j - G_j)\cos\left(\frac{\pi}{2} - \tau_j\right) r\sin(\theta - \theta_j)$$

$$- (W_j - G_j)\sin\left(\frac{\pi}{2} - \tau_j\right)[r_j - r\cos(\theta - \theta_j)]$$

$$= (W_j - G_j)[r\cos(\theta - \psi_j) - r_j\cos(\theta_j - \psi_j)].$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Adding up the contributions of nodes 1 to $i-1$, we obtain:

$$M(\theta) = \sum_{j=1}^{i-1} (W_j - G_j)[r\cos(\theta - \psi_j) - r_j \cos(\theta_j - \psi_j)],$$

$$\theta_{i-1} \le \theta < \theta_i, \quad i = 2, \ldots, N. \tag{17}$$

Note that H, V, P, and M are piecewise smooth functions of θ.

The loads must satisfy the equilibrium conditions of the ring:

$$H(\theta_N) = V(\theta_N) = M(\theta_N) = 0 \quad (18)\ (19)\ (20)$$

which are obtained from equations (15) and (17) by summing the respective right-hand-side expressions from $j = 1$ to N. Equations (18), (19), and (20) are the needed conditions to handle the additional unknowns c_r, c_θ, c_ψ.

2.4 Conditions for Contact and Separation. Let us choose a coordinate system fixed to the engine block, in which the polar angle ϕ is measured from the minor thrust side, Fig. 4. The bore distortion and the bore temperature rise are given as functions of ϕ. The angular coordinate θ of the ring (and likewise, θ_o of the free ring) is measured from the middle of the free gap. The relation between the two polar angles is:

$$\phi = \theta_s + \theta \tag{21}$$

where θ_s specifies the angular position of the free gap relative to the bore.

An effective bore shape is defined as:

$$r_b = R_b + \delta_b - \left(\gamma + \frac{a}{2}\right) \tag{22}$$

where R_b is the nominal bore radius, δ_b is the bore distortion, and the terms in parentheses take account of the finite wall thickness of the ring. A separation function is then defined as:

$$v = r_b - r. \tag{23}$$

Recall that r must be given by the solution of equations (12), (13), and (9′), which depend on the contact loads. Therefore, the right-hand side of equation (23) may be viewed as a nonlinear operator, Ω, on the contact loads. Clearly, v must be non-negative. When $v = 0$, equation (23) is to be solved for the contact loads, which should also be non-negative.

To proceed further, we discretize equation (23). The discrete problem is to find v_i and W_i, $i = 1, 2, \ldots, N$, such that they satisfy the following conditions:

$$v_i = \Omega[W_1, W_2, \ldots, W_N]; \tag{24}$$

$$\left.\begin{aligned} v_i &= 0, \quad W_i \ge 0, \quad \text{for } i \in C; \\ v_i &< 0, \quad W_i = G_i, \quad \text{for } i \in S. \end{aligned}\right\} \tag{25}$$

This is a complementarity problem involving bounds (Sargent, 1978). Algorithms are available for the solution of this type of problem.

2.5 Question of Solution Uniqueness. The mathematical problem can therefore be summarized as: Find a partition of the total number of nodes into the set of contact nodes (C) and the set of separation nodes (S), such that the W_i's obtained from solving $v_i = 0$, $i \in C$ are non-negative and such that the W_i's ensure $v_i > 0$, $i \in S$. The radial coordinate r of the ring, which is needed in the expression for v, is to be obtained from integrating the system of equations (12), (13), and (9′); and will depend on three initial values c_r c_θ c_ψ, whose determination requires the equilibrium conditions (18), (19), and (20).

It is conceived, however, that the problem stated above may possess multiplicity of solutions. Suppose the bore distortion consists of two symmetric lobes. A ring of any free shape would have at least two equilibrium positions, if it has any, when inserted inside the bore. If, further, the ring is symmetric (with respect to an axis passing through the foot and the middle of the gap), it would have at least four equilibrium positions. For general bore distortion and free ring shape, it is hard to see how many solutions there may be, and one will likely encounter difficulties in effecting a solution.

On the other hand, if the angular position of the ring is fixed, the question of solution uniqueness may be avoidable. Indeed, for a circular (undistorted) bore, the ring is indifferent to its angular position relative to the bore, and the latter must be specified. In this case the number of independent equilibrium conditions is also reduced to two, because the contact loads act on a circle and become concurrent. Hence, the number of unknowns and number of equations are again compatible. This case suggests that we may specify the ring gap position if we treat the ring deflection as small perturbation from a circle. This we shall carry out in the next section.

It should be mentioned that the small perturbation analysis does not unduly restrict the usefulness of the solution. In usual practice the ring deflection is two orders of magnitude smaller than the size of the bore.

3 Method of Solution

3.1 Linearized Equations. Let us assume that the free ring shape is nearly a circle and that ring deflection is small compared with the radius of the circle. We take this circle to be the one passing through the neutral axis of the free ring at $\theta_o = \pi$. See Fig. 4. With respect to this reference circle whose radius to be denoted as R_m, let us introduce small perturbations:

$$\left.\begin{aligned} r &= R_m - u, & |u| &\ll R_m \\ \tau &= \frac{\pi}{2} + \omega, & |\omega| &\ll \frac{\pi}{2} \\ R_{no} &= R_m - \rho, & |\rho| &\ll R_m. \end{aligned}\right\} \tag{26}$$

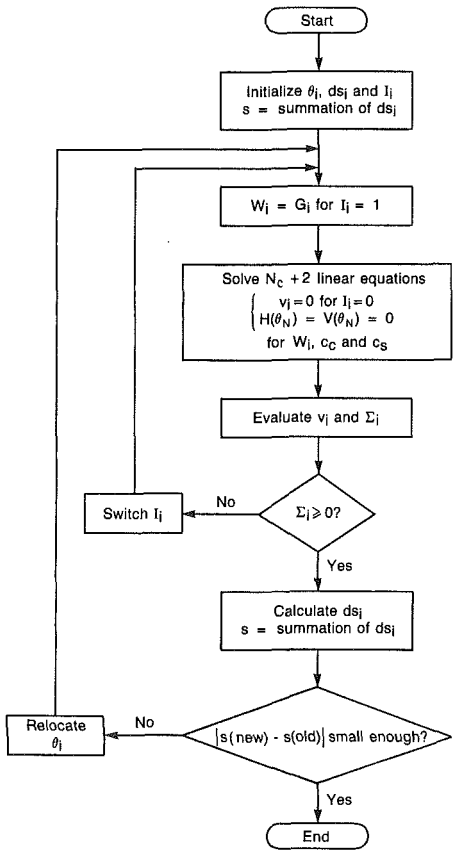


Fig. 5 Solution procedure

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Consistent with the small deflection restriction, we also assume:

$$\left|\frac{P}{EA}\right| \ll 1, \quad \left|\frac{M}{E\gamma A}\right| \ll 1, \text{ and } |\alpha T| \ll 1. \tag{27}$$

The ring cross-section may vary. But as long as its variation is mild, γ and A need not be expanded. This is because these quantities are only associated with the small terms listed in equation (27).

Substituting equation (26) and (27) into (12), (13), and (9′), and neglecting higher-order small terms, we obtain a corresponding system of linear equations:

$$\frac{du}{ds} = \omega \tag{28}$$

$$\frac{d\theta}{ds} = \frac{1}{R_m}\left(1 + \frac{u}{R_m}\right) \tag{29}$$

$$\frac{d}{ds}(\theta+\omega) = \frac{1}{R_m}\left\{1 + \frac{M}{E\gamma A} + \frac{\alpha}{\gamma A}\int_A Tz\,dA - \frac{P}{EA} - \frac{\alpha}{A}\int_A T\,dA + \frac{\rho}{R_m}\right\}. \tag{30}$$

Eliminating ω and s among equations (28)–(30), we arrive at a single equation governing the small deflection u, with θ as the independent variable:

$$\frac{d^2u}{d\theta^2} + u = R_m\left\{\frac{M}{E\gamma A} + \frac{\alpha}{\gamma A}\int_A Tz\,dA - \frac{P}{EA} - \frac{\alpha}{A}\int_A T\,dA\right\} + \rho \equiv RHS. \tag{31}$$

The expressions for the moment and forces on a ring cross-section can be similarly expanded. Since these quantities are themselves small, only 0(1) terms in the expansions need be retained. Hence, from equations (15)–(17) we have:

$$H(\theta) = \sum_{j=1}^{i-1}(W_j - G_j)\cos\theta_j \tag{32}$$

$$V(\theta) = \sum_{j=1}^{i-1}(W_j - G_j)\sin\theta_j \tag{33}$$

$$P(\theta) = -H\sin\theta + V\cos\theta = -\sum_{j=1}^{i-1}(W_j - G_j)\sin(\theta - \theta_j) \tag{34}$$

$$M(\theta) = \sum_{j=1}^{i-1}(W_j - G_j)R_m\sin(\theta - \theta_j). \tag{35}$$

These expressions hold for $\theta_{i-1} \le \theta < \theta_i$, $i = 2, \ldots, N$.

3.2 Ring Deflection in Terms of the Angular Coordinate. Equation (31) can be solved by the method of variation of parameters. The result is:

$$u(\theta) = \left[c_s + \int_{\theta_1}^{\theta} RHS\cos\theta'\,d\theta'\right]\sin\theta + \left[c_c - \int_{\theta_1}^{\theta} RHS\sin\theta'\,d\theta'\right]\cos\theta, \quad \theta_1 \le \theta \le \theta_N \tag{36}$$

where c_c and c_s are two integration constants. If the free ring shape, the cross-sectional properties γ and A, and the temperature field are given in closed form, the integrals in equation (36) may also be evaluated in closed form. If these quantities are given in the form of discrete data, then the integrals must be evaluated numerically. Among the terms in RHS, M and P involve the unknowns $(W_i - G_i)$. The integration of the M and P terms yields the coefficients of these unknowns. For the case of a uniform ring cross-section, these coefficients can be reduced to simple form. After somewhat lengthy manipulations, the resulting equation (36) in discrete form reads:

$$\left.\begin{aligned} u_i = {} & \left\{c_s + \int_{\theta_1}^{\theta_i}\left(\frac{\sigma R_m}{\gamma A}\int_A Tz\,dA - \frac{\alpha R_m}{A}\int_A T\,dA + \rho\right)\cos\theta'\,d\theta'\right\}\sin\theta_i + \left\{c_c - \int_{\theta_1}^{\theta_i}\left(\frac{\alpha R_m}{\gamma A}\int_A Tz\,dA - \frac{\alpha R_m}{A}\int_A T\,dA + \rho\right)\sin\theta'\,d\theta'\right\}\cos\theta_i \\ & + \left(\frac{R_m^2}{E\gamma A} + \frac{R_m}{EA}\right)\frac{1}{2}\sum_{j=1}^{i}(W_j - G_j)[\sin(\theta_i - \theta_j) - (\theta_i - \theta_j)\cos(\theta_i - \theta_j)], \\ & i = 1, \ldots, N. \end{aligned}\right\} \tag{37}$$

Notice that the three unknown initial values c_r, c_θ, and c_ψ are now replaced by the two integration constants c_c and c_s. This is because, since the ring deflection is treated as small

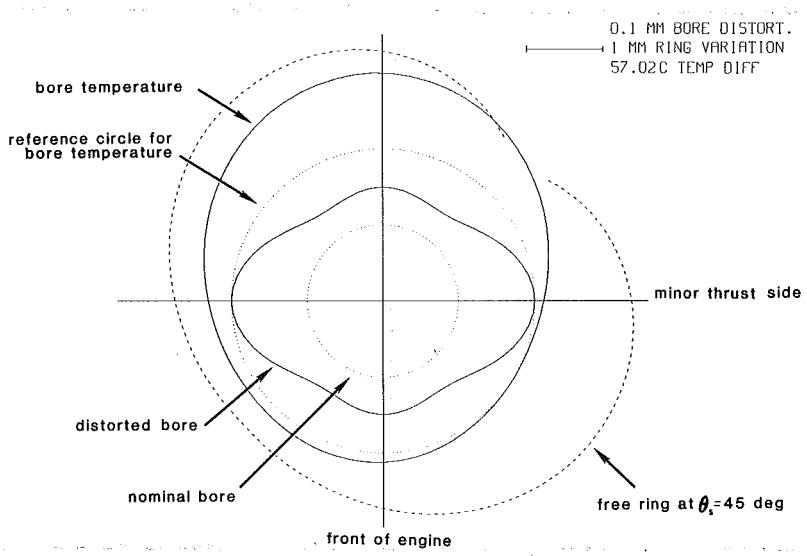


Fig. 6 Free ring shape, bore distortion, and bore temperature variation

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Numerical Values Used in Computation

Ring material values:
$E = 89.6$ GPa
$\alpha = 11.4 \times 10^{-6}/°C$

Ring dimensions:
end clearance = 3.631 mm
free gap = 12.774 mm
$R_b = 44.45$ mm
$a = 4.07$ mm
$b = 1.966$ mm
$\delta_r = 1.0$ mm (round ring)

Bore distortion:
$\delta_{bo} = 0.06$ mm
$\delta_{b2} = 0.025$ mm
$\delta_{b4} = 0.015$ mm

Temperature rise:
$T_p = 240°C$
$\delta_{To} = 140°C$
$\delta_{T1} = 25°C$
$\delta_{T2} = -5°C$
$\delta_{T3} = 0°C$
$\delta_{T4} = -10°C$
$T_b = 150°C$ (uniform bore temperature)

Gas pressure values
0, and 3.45 MPa

Free gap positions:
0 deg, 45 deg, and 90 deg

perturbation from a *circle*, the deflected ring is indifferent to its angular position relative to the circle. Hence, this relative angular position must now be *specified* for the problem, rather than *determined* by it.

3.3 Equilibrium Conditions. Substituting the expansions (26) into equations (18), (19), and (20) and neglecting small terms we obtain:

$$H(\theta_N) = \sum_{j=1}^{N} (W_j - G_j)\cos\theta_j = 0 \tag{38}$$

$$V(\theta_N) = \sum_{j=1}^{N} (W_j - G_j)\sin\theta_j = 0 \tag{39}$$

$$M(\theta_N) = \sum_{j=1}^{N} (W_j - G_j)R_m\sin(\theta_N - \theta_j) = 0. \tag{40}$$

Clearly, equations (38) and (39) imply the validity of equation (40). Hence, only the first two are independent equilibrium conditions. This is the result of the small deflection approximation, by which the loads act on a circle and are concurrent at its center.

3.4 Locating the Angular Coordinates of Nodes. It re-

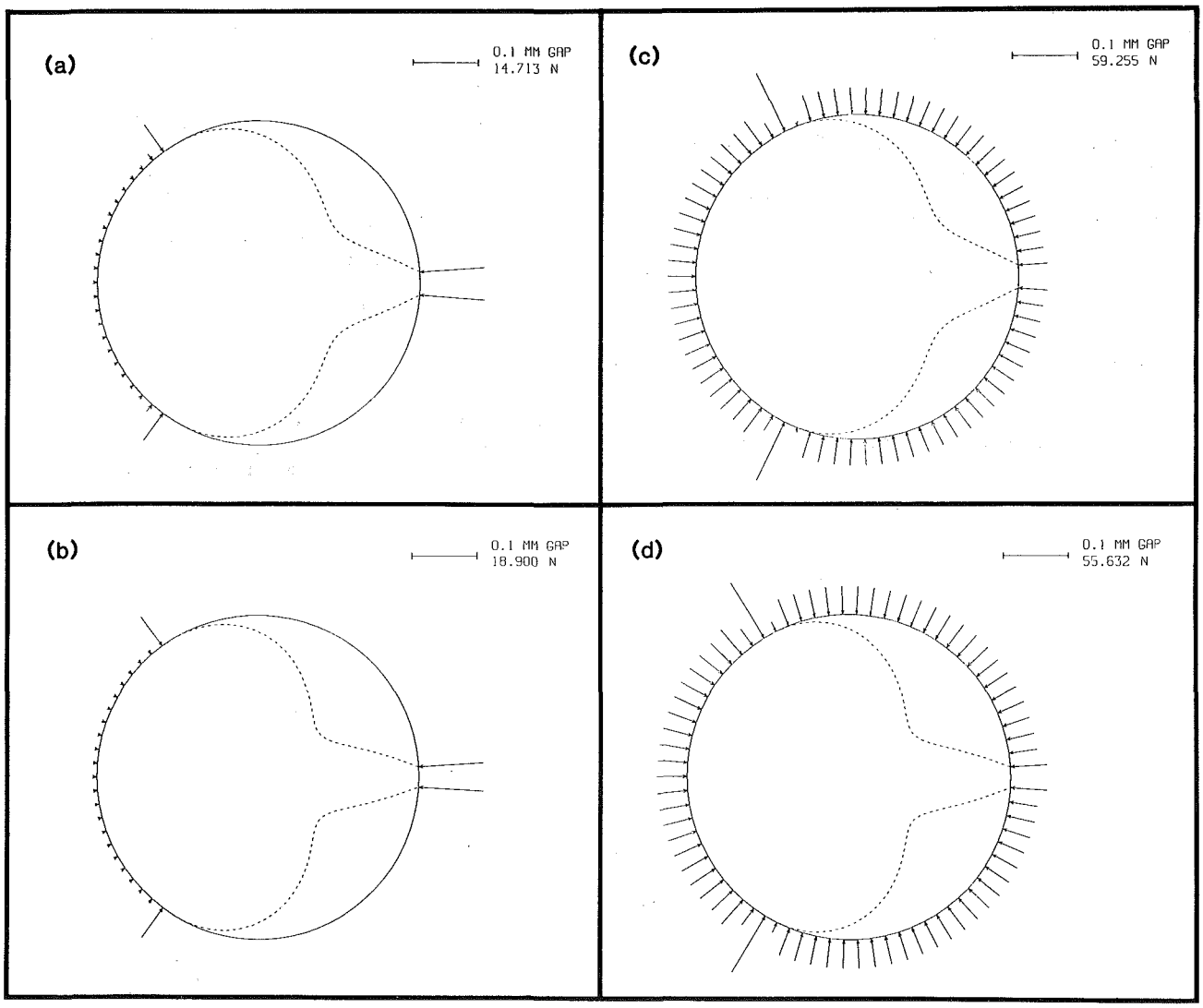


Fig. 7 Contact pattern and loads—round ring with undistorted bore

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

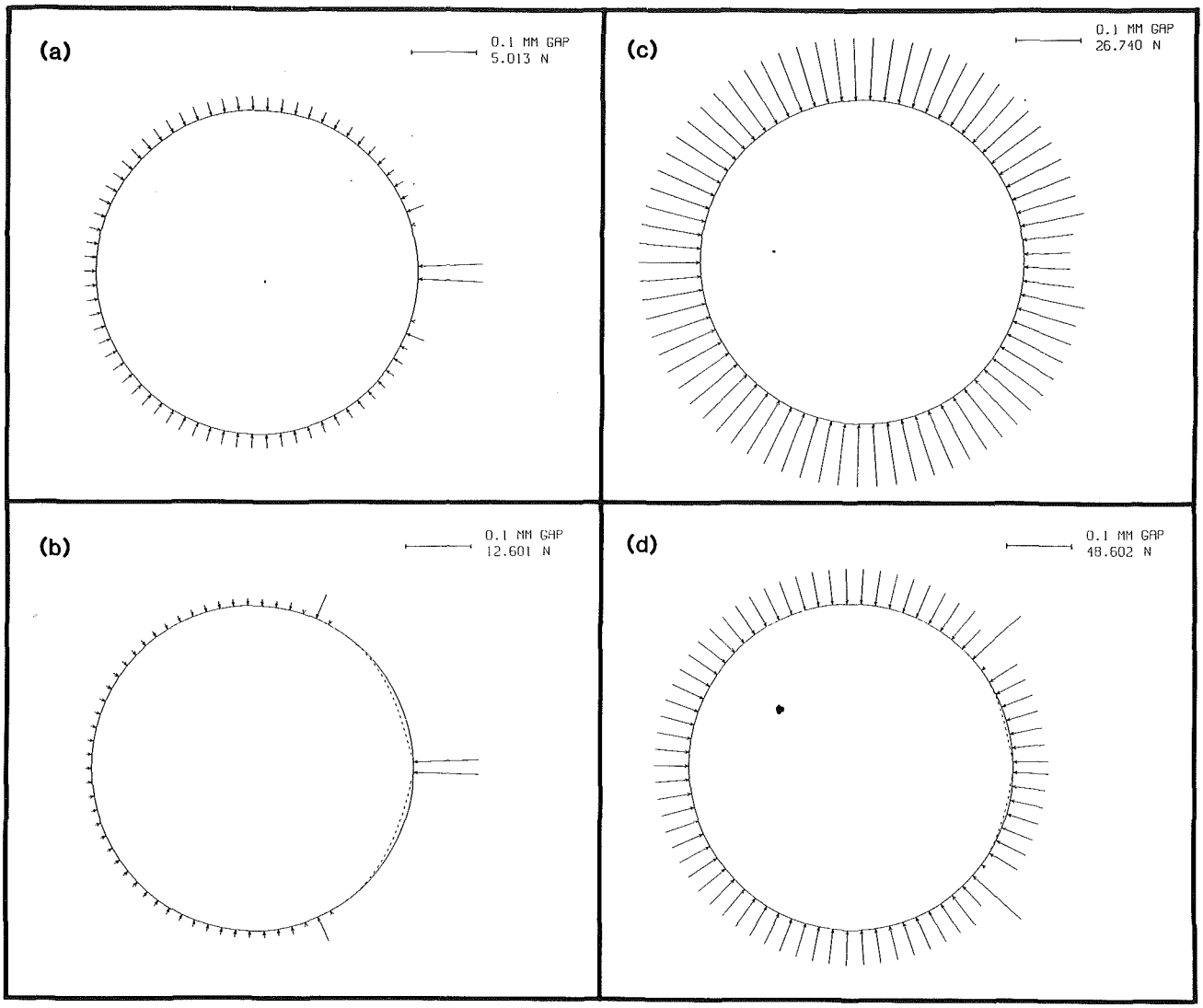


Fig. 8 **Contact pattern and loads—handbook ring with undistorted bore**

mains to determine θ_i based on the length of the deflected ring. We choose to fix the middle node of the deflected ring at $\theta = \pi$. The other angular coordinates are then located by the relation:

$$d\theta = \frac{ds}{\sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}} \tag{41}$$

where ds is given by equation (11). Since these angular coordinates depend on the radial coordinates and loads, whose determination in turn depends on the angular coordinates, iteration is needed.

3.5 Solution Procedure. The solution scheme consists of two iteration loops. The outer one corrects for the angular coordinates of the nodes. The inner one is the complementarity loop that searches for the contact pattern.

The complementarity loop is implemented by using the following algorithm: An index set I_i and an inspection set Σ_i are first constructed:

$$\left.\begin{aligned} I_i = 0, \quad \Sigma_i = v_i + W_i, \quad & \text{for } i \in C \\ I_i = 1, \quad \Sigma_i = v_i + (W_i - G_i), \quad & \text{for } i \in S. \end{aligned}\right\} \tag{42}$$

We begin with a particular choice of the index set. Based on this choice the W_i's and c_c, c_s are solved and the v_i's evaluated. Then we inspect the Σ_i values simultaneously from both ends of the ring. We find from each end the first node where the Σ_i value is negative, and move the index value there from its original category to the opposite category. In this way a new index set is obtained, which differs from the previous one by no more than two index values. The same procedure is carried out with this new index set, and the cycles continue until all the Σ_i values are non-negative. The complementarity loop is then completed.

It should be mentioned that the algorithm used is a heuristic modification of the direct algorithm developed by Habetler and Kostreva (1978), in which at most one index value may be changed in each cycle. The author does not have a convergence proof of the modified version, but its application to the present problem has been successful.

Within each complementarity cycle, a system of equations need be solved to determine the W_i's and the additional unknowns c_c, c_s. Suppose the index set for this cycle contains N_c ($\leq N$) zeroes. Then, the system is of size $N_c + 2$, consisting of equations (38) and (39) and N_c equations from (25). A subroutine, LEQT1F, of the IMSL Library is used to solve this system.

Figure 5 summarizes the solution procedure.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

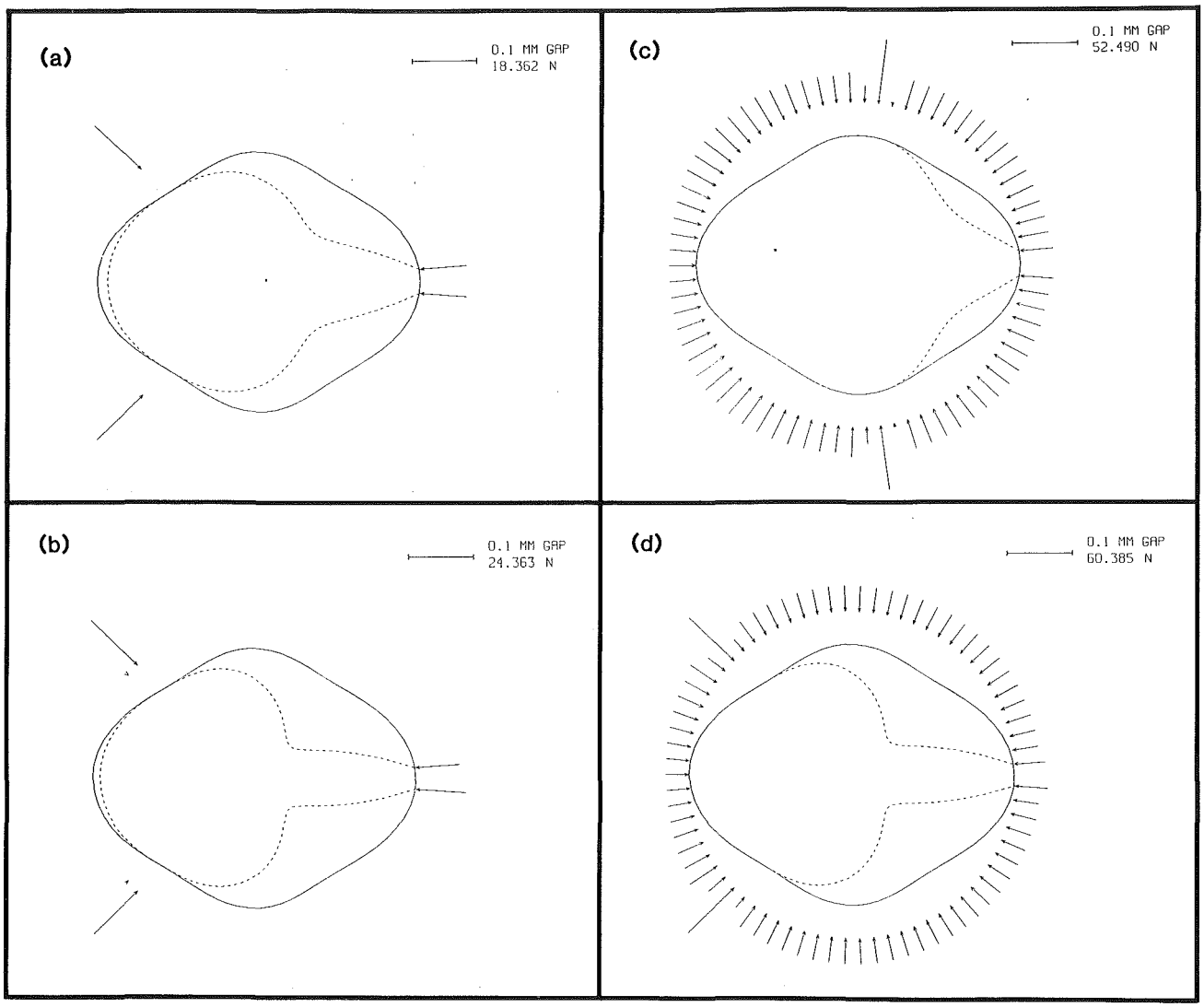


Fig. 9 Contact pattern and loads—round ring with distorted bore, θ_s = 0 deg

4 Input Data for Computation

A ring of rectangular cross-section is used for the computation. Two free ring shapes (given by the radial coordinate of the ring face, r_f) are considered. One is circular with a radius slightly larger (by an amount δ_r) than that of the nominal bore:

$$r_f = R_b + \delta_r. \tag{43}$$

This shape will hereafter be referred to as the *round ring*. The other shape is taken from a ring designer's handbook (Koppers Company, 1975):

$$r_f = R_b + \frac{g}{3\pi}\left(1 + \frac{1}{2}(\pi - \theta_o)\sin\theta_o\right) \tag{44}$$

where g is free gap less end clearance. This shape will be referred to as the *handbook ring*. The dimensions and material values of the rings are listed in Table 1.

Bore distortion is mainly caused by bolt-up stresses and uneven thermal expansion. It can be approximated by a three-term empirical formula:

$$\delta_b = \delta_{bo} + \delta_{b2}\cos 2\phi + \delta_{b4}\cos 4\phi \tag{45}$$

where δ_{b0} and δ_{b2} represent, respectively, the uniform and uneven parts of the thermal expansion, and δ_{b4} represents the bolt-up effect. Values for these coefficients near the top dead center (TDC) are listed in Table 1.

The bore temperature rise can be represented by an empirical formula:

$$T_b = \delta_{To} + \delta_{T1}\sin\phi + \delta_{T2}\cos\phi + \delta_{T3}\sin 2\phi + \delta_{T4}\cos 2\phi. \tag{46}$$

Values for the coefficients in this formula, near the TDC, and the corresponding piston temperature rise are listed in Table 1. The case of equation (46) will be known as one of *non-uniform bore temperature*. In addition, a case of *uniform bore temperature* is also used in the computation, where there is a uniform temperature drop of 90°C between the piston and the bore.

The temperature field in the ring is approximated by allowing a 20 percent temperature drop across the piston-ring gap, a 20 percent drop across the ring-bore gap, and a linear variation within the ring:

$$T = 0.2T_p + 0.8T_b + \frac{0.6(T_p - T_b)}{a}\left(\frac{a}{2} + \gamma - z\right). \tag{47}$$

With equation (47), the integrals in equation (37) involving the ring temperature are reduced to:

$$\begin{aligned} \frac{1}{A}\int_A T dA &= \frac{1}{2}(T_p + T_b) \\ \frac{1}{\gamma A}\int_A T z dA &= \frac{1}{2}(T_p + T_b) - \frac{a}{20\gamma}(T_p - T_b). \end{aligned} \tag{48}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

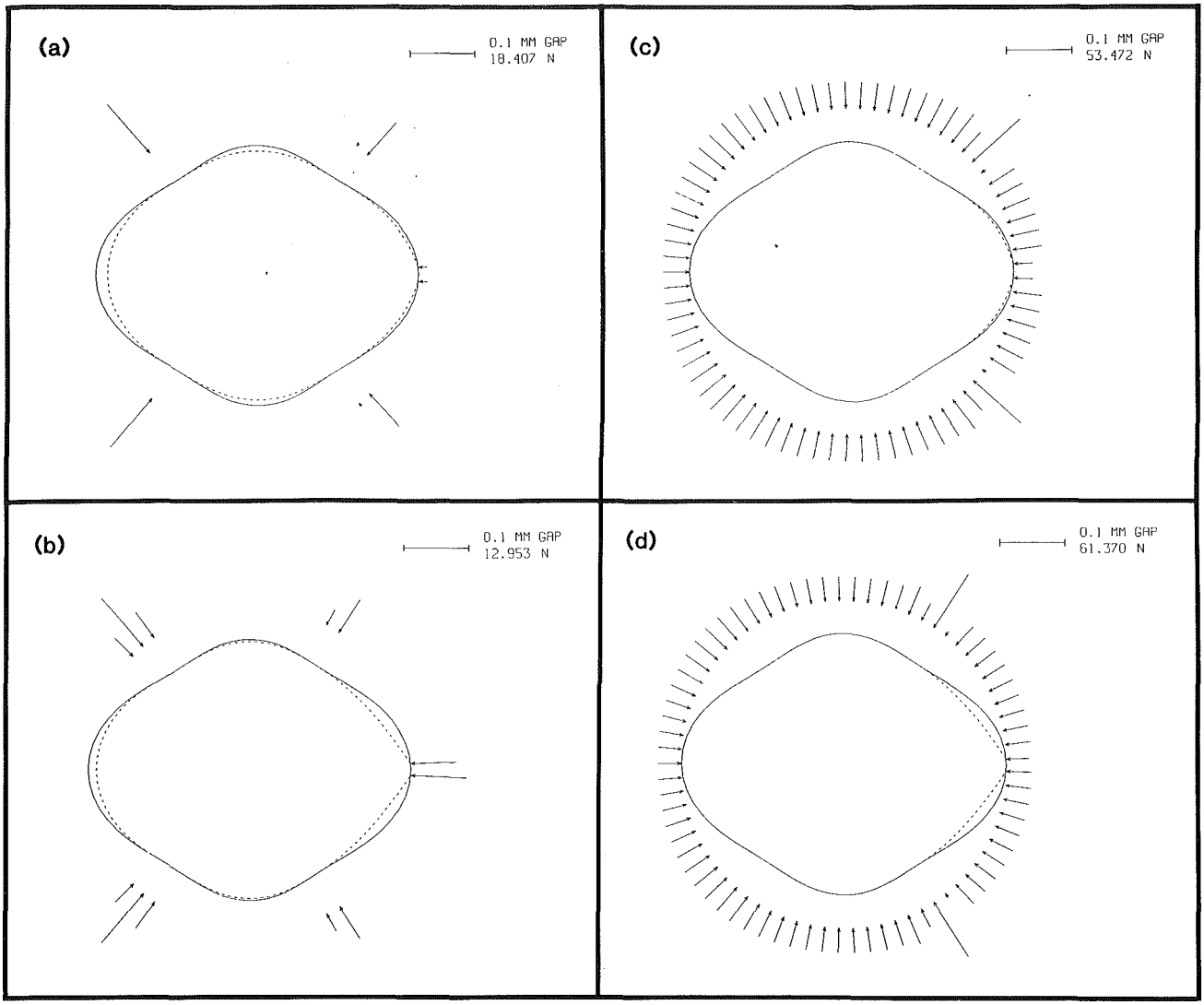


Fig. 10 Contact pattern and loads—handbook ring with distorted bore, $\theta_s = 0$ deg

The gas pressure values and free gap positions used in the computation are listed in Table 1.

Figure 6 shows the handbook ring at $\theta_s = 45$ deg, the distorted bore as given by equations (45), and the bore temperature variation as given by equations (46).

5 Results and Discussion

5.1 Grid Size and Convergence. The effect of grid size is investigated with two test cases (round and handbook rings with undistorted bore, no thermal deformation of the ring and no gas pressure) where uniform contact regions are known to exist. The grid size $\Delta\theta_o$ is varied from 10 deg (the coarsest grid) to 2 deg (the finest). It is found that the computed contact load values become stabilized for $\Delta\theta_o \leq 5$ deg. Hence, a grid size of 5 deg ($N = 73$) is chosen for the rest of the computation.

The outer loop locating the angular coordinates of the nodes converges rapidly. Four iterations are enough to ensure convergence which requires that two consecutively computed ring lengths differ by less than 10^{-8} mm. The complementarity algorithm also converges nicely, and in a regular pattern. Only in one of the computed cases a filter (Kostreva, 1984) of $v_i \geq -0.8\ \mu$m, $i \in S$ is needed to halt the oscillation of index values near the solution. In all other cases the straightforward criterion $v_i \geq 0$ works successfully.

5.2 Organization of Figures. The computed ring-bore separation distance (v) and contact loads (W_i) are presented in polar plots, first for the undistorted bore (Figs. 7 and 8), then for the distorted bore (Figs. 9–14). The figures for the distorted bore are divided into three groups, viz. $\theta_s = 0$ deg (Figs. 9 and 10), 45 deg (Figs. 11 and 12), and 90 deg (Figs. 13 and 14). The odd-numbered figures are for the round ring; the even-numbered figures are for the handbook ring. Each set of Figs. 7–14 consist of four cases. They are:

(a) no gas pressure, no thermal deformation of the ring,
(b) no gas pressure, with thermal deformation of the ring,
(c) with gas pressure, no thermal deformation of the ring, and
(d) with gas pressure, with thermal deformation of the ring.

Uniform bore temperature is used in the case of the undistorted bore, and nonuniform bore temperature is used in the distorted bore. In these figures, v is given by the radial distance between the dashed curve (the ring) and the solid curve (the bore), and W_i by the inward pointing arrows. It should be noticed that the gas loads acting on the ring back are not shown even when they are present. Note also that the scale for v is the same in all the figures, whereas that for W_i varies from figure to figure. Thus, comparing contact patterns is straightforward, while interpreting load magnitudes in different cases needs some care.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

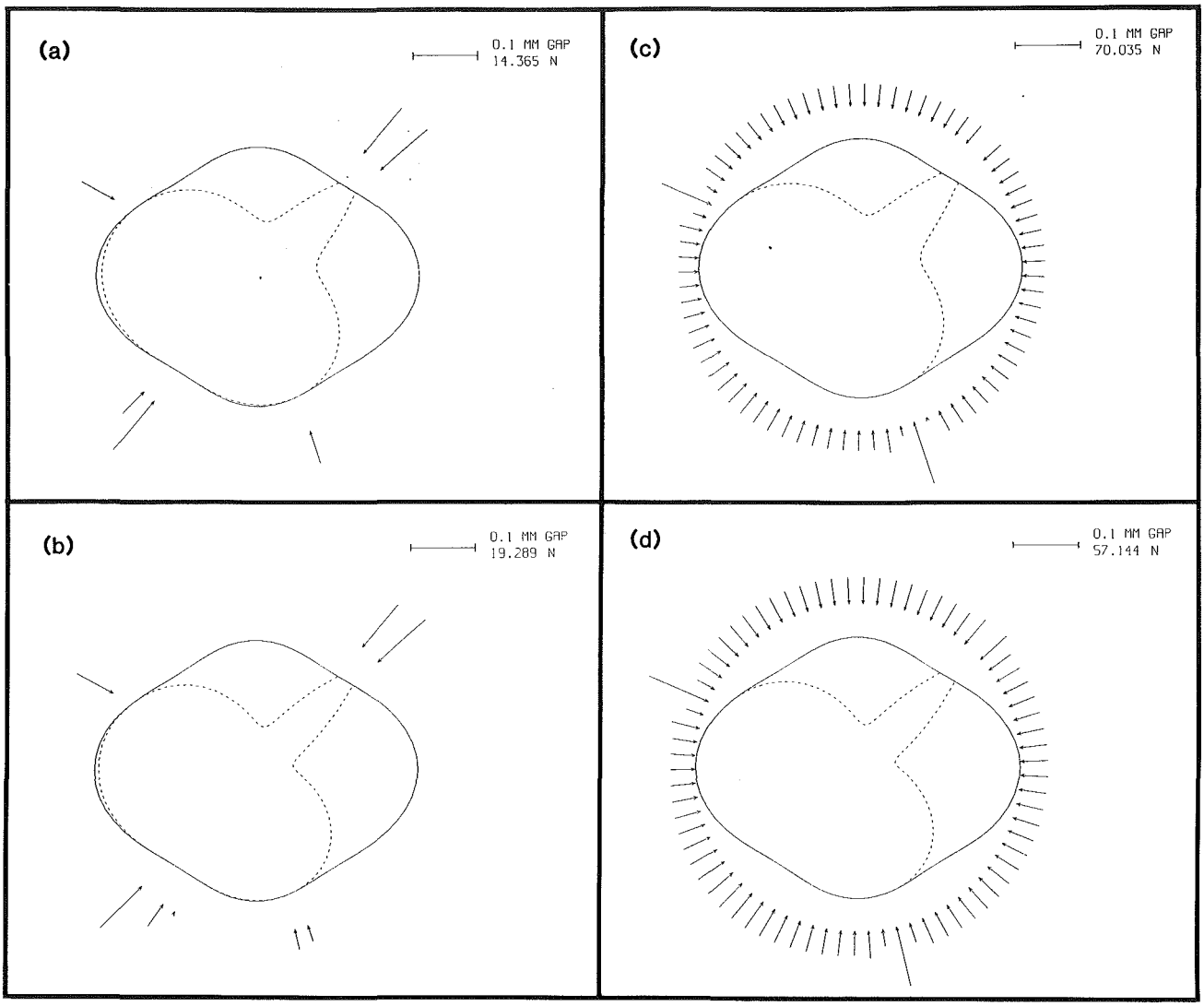


Fig. 11 Contact pattern and loads—round ring with distorted bore, θ_s = 45 deg

5.3 Comparison With Analytical Solution. The case shown in Fig. 7(*a*) has an analytical solution (Feodosyev, 1977), as previously mentioned. That solution predicts a uniform contact region (with distributed load q) opposite the ring gap, a concentrated load Q_2 at the edge of the uniform contact region, and a concentrated load Q_1 at the ring end. These quantities in our notation are:

$$q = \frac{EI\delta_r}{R_b^4}, \quad Q_2 = 0.64\frac{EI\delta_r}{R_b^3}, \quad Q_1 = 1.19\frac{EI\delta_r}{R_b^3} \tag{49}$$

where $I = ba^3/12$ is the moment of inertia of the cross-section. Table 2 compares our numerical solution with this solution.[2] The comparison is considered reasonable because (1) the numerical solution contains various discretization errors, (2) the analytical solution neglects the curvature effect on the location of the neutral axis, and (3) the analytical solution assumes vanishing end clearance.

5.4 Undistorted Bore. For the case shown in Fig. 7(*b*), a uniform bore temperature of 150°C and a piston temperature of 240°C are used. The thermal deformation of the ring results in higher loads (Q_1 = 18.9N, q = 0.376N/mm) and larger separation distance (v_{max} = 0.134 mm) that in Fig. 7(*a*), where v_{max} is 0.105 mm. But the extent of the uniform contact region remains almost the same. Figures 7(*c*,*d*) show the effect of gas pressure, which increases the uniform contact region by 15 deg in Fig. 7(*c*) and 10 deg in Fig. 7(*d*) but reduces the v_{max} values only slightly (0.098 mm in Fig. 7(*c*), 0.131 mm in Fig. 7(*d*)). The load increase in these figures mainly reflects the magnitude of gas pressure. The loads plotted in the separation region show the magnitudes of the discrete gas loads resulting from the gas pressure.

The state of contact of the handbook ring is greatly improved, as shown in Figs. 8. In Fig. 8(*a*) uniform contact is achieved almost everywhere. Though there is a separation region extending from the ring end to θ = 26.9 deg, the v_{max} value is only 0.1 μm. The distributed contact load in this case is 0.284N/mm at θ = 180 deg.[3] With thermal deformation of the ring, Fig. 8(*b*), the separation region increases to 66.5 deg,

[2] Q_2 for the numerical solution is taken to be the largest load closest to the edge of the uniform contact region. It contains both the concentrated load and the contribution of the distributed load.

[3] The uniform contact solution by Swift (1947) gives the contact load as, in our notation, $q = gEI/(3\pi R_b^4)$, or 0.246 N/mm.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

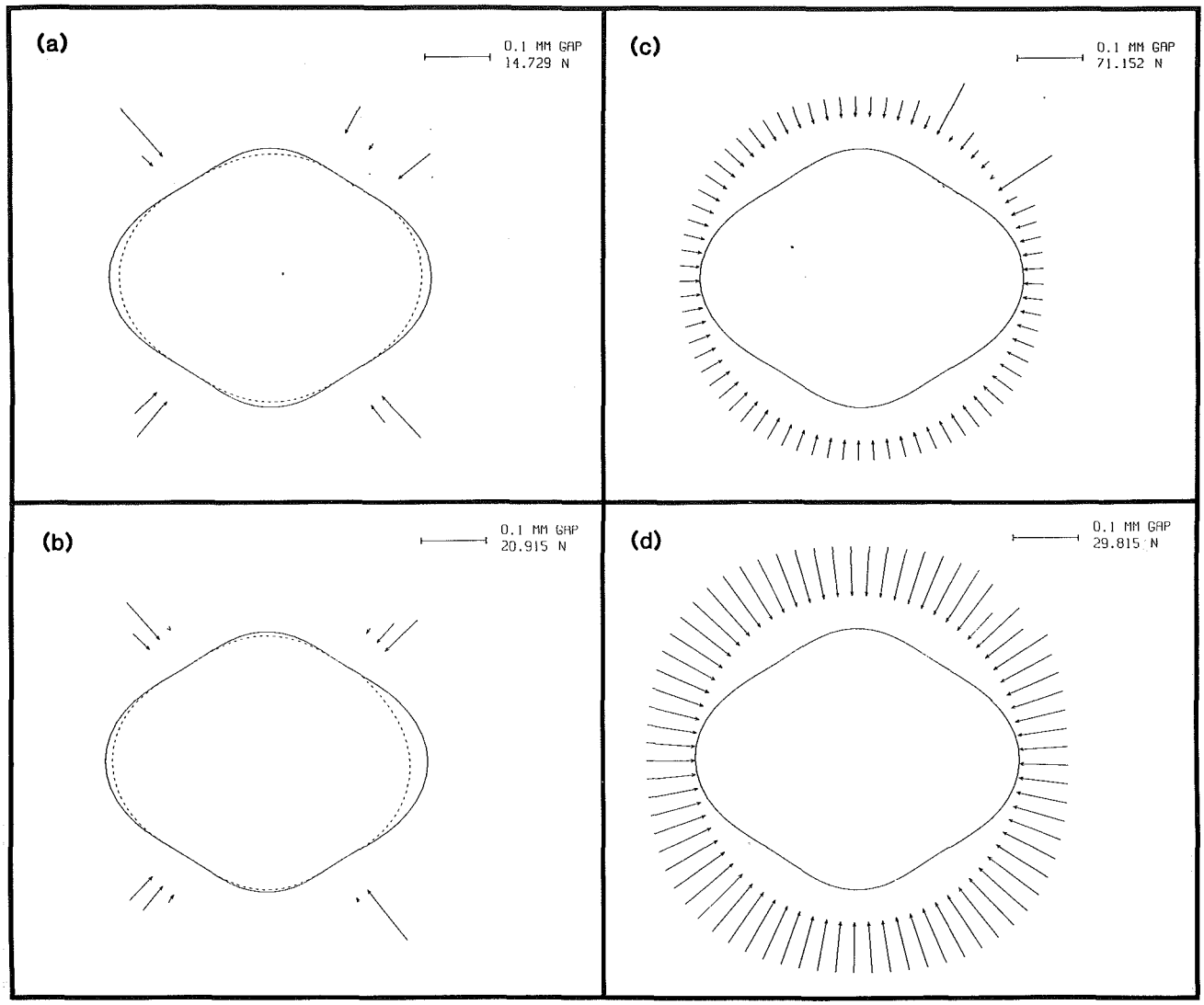


Fig. 12 Contact pattern and loads—handbook ring with distorted bore, θ_s = 45 deg

Table 2 Comparison With Analytical Solution

	Numerical Solution	Analytical Solution
Q_1, N	14.7	13.4
Q_2, N	7.28	7.21
q, N/mm	0.292	0.254
θ at Q_2, deg	126.3	122.6

the v_{max} value to 8.2 μm, and the distributed contact load to 0.369N/mm. With the help of gas pressure, Figs. 8(c,d), the small separation region in Fig. 8(a) disappears, and the separation region in Fig. 8(b) is greatly reduced.

5.5 Distorted Bore. Contact patterns change drastically when bore distortion is present. Let us compare Figs. 7(a,b) with Figs. 9(a,b). The uniform contact regions present in the former cases totally disappear in the latter. The contact loads in the latter cases become discrete in nature, though the load magnitudes and v_{max} values remain of the same order. Gas pressure still improves ring conformity, but its effectiveness is compromised when combined with temperature rise. Thus, in the cases of Figs. 9(a,c), gas pressure brings about a major improvement; whereas in the cases of Figs. 9(b,d), it is only effective in eliminating the small separation region opposite the ring gap. Notice that the contact patterns in Figs. 9(b,d) are not symmetric, because the temperature rise is not symmetric.

The strong influence of bore distortion becomes more noticeable when one compares Figs. 8(a,b) with Figs. 10(a,b). In addition to the change in contact nature, the separation distance also increases manyfold. Thus, v_{max} is 16.9 μm in Fig. 10(a) (compared to 0.1 μm in Fig. 8(a)), and is 18.7 μm in Fig. 10(b) (compared to 8.2 μm in Fig. 8(b)). Gas pressure is capable of eliminating most separation regions in the cases of Figs. 10(a,c); but in the cases of Figs. 10(b,d), it can neither reduce the extent of the separation region nor the separation distance near the ring end.

5.6 Position of Ring Gap. Figures 11–14 display the influence of ring gap position. By carefully comparing these figures with Figs. 9 and 10, one finds that the previously described contact features remain qualitatively the same. But quantitative differences exist, which can be either improvement or deterioration. For the round ring, θ_s = 45 deg appears detrimental. For instance, v_{max} is 0.19 mm in Fig. 11(b), a 50 percent increase over that in Fig. 9(b). On the other hand, the handbook ring adapts itself well at θ_s = 90 deg. The state of contact in Figs. 14 is generally better than that in Figs. 10.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

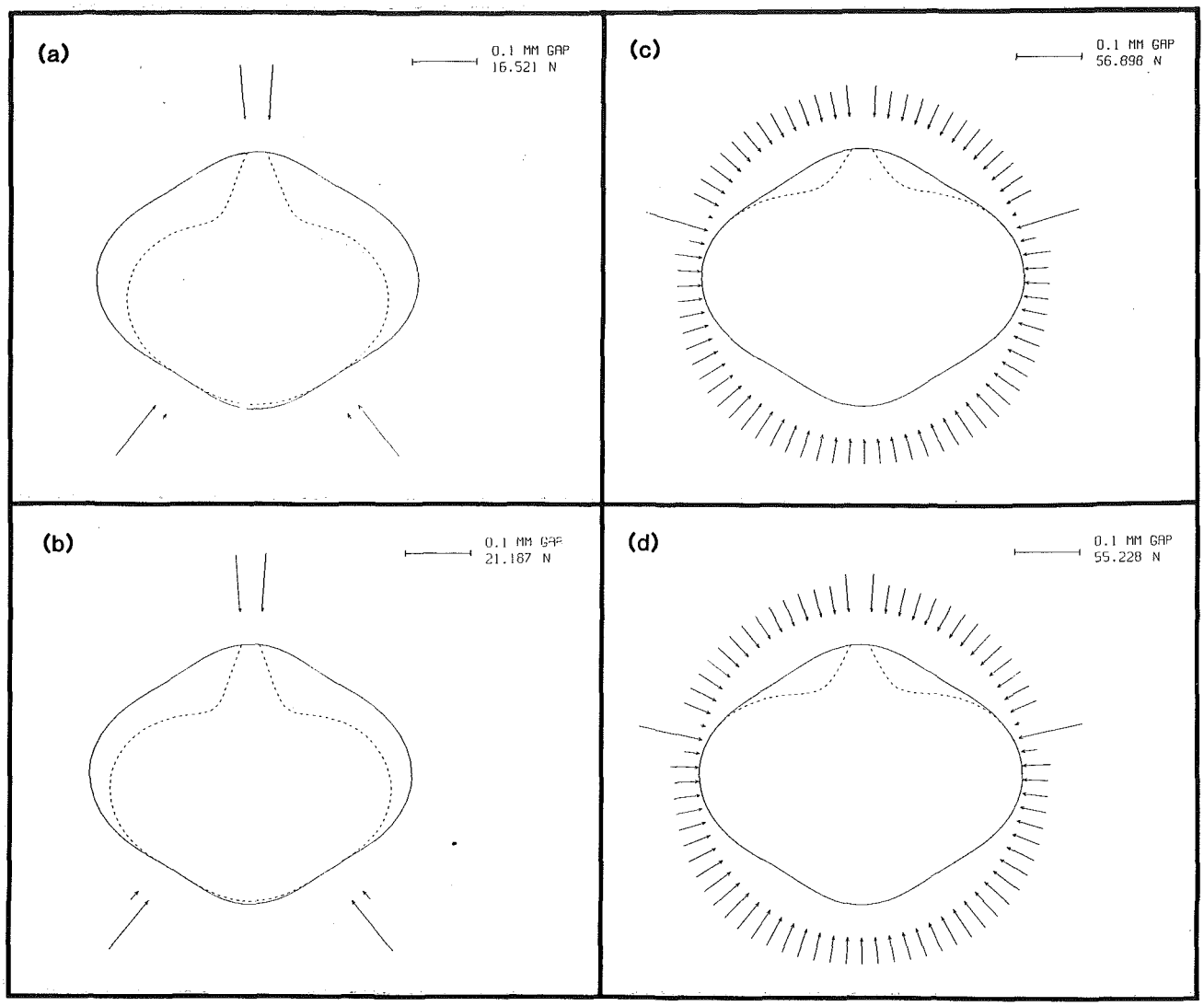


Fig. 13 Contact pattern and loads—round ring with distorted bore, θ_s = 90 deg

6 Closure

The contact between piston ring and cylinder bore is analyzed, taking into account arbitrary bore shape, arbitrary free ring shape, thermal deformation of the ring, and the effect of gas pressure. A solution scheme is devised to implement computation, and is validated by comparing results in special cases with known analytical solutions.

Results of computation based on dimensions, material values, and operating parameters commonly found in automotive engines have shown the following: (1) The handbook ring has the correct shape for the undistorted bore. But when bore distortion is present, nonuniform contact becomes a rule rather than an exception, and is accompanied with manyfold increase in magnitude of the contact loads. (2) Thermal deformation of the ring can both alter the contact pattern and increase the contact loads significantly. This thermal effect makes it more difficult to achieve uniform contact. (3) Gas pressure improves ring conformity, but it cannot totally compensate for the adverse state of contact caused by bore distortion and thermal deformation of the ring. (4) Different ring gap positions can either improve or deteriorate the state of contact.

The work has several possible applications: (1) The information on contact pattern and loads can serve as the basis of refined lubrication and friction analyses. (2) The information on ring-bore separation can be used to study gas blowby and oil consumption. (3) The solution scheme can be used as a design tool for the free ring shape to achieve better ring-bore contact in specific bore distortion and operating conditions.

Acknowledgment

The author is indebted to Dr. Chin-Hsiu Li of General Motors Research Laboratories for providing the bore distortion and temperature data. Many discussions were held with Dr. Li during the initial phase of this study, and these discussions proved most stimulating and helpful.

References

Boley, B. A., and Weiner, J. H., 1960, *Theory of Thermal Stresses*, John Wiley and Sons, New York, Chapter 11.

Dragoni, E., and Strozzi, A., 1986, "Analysis of a Split Ring Inserted into a Circular Housing," *Journal of Strain Analysis*, Vol. 21, pp. 59–70.

Feodosyev, V. I., 1977, *Selected Problems and Questions in Strength of Materials*, Mir Publishers, Moscow, pp. 168–170.

Frisch-Fay, R., 1962, *Flexible Bars*, Butterworths, Chapter 4.

Habetler, G. J., and Kostreva, M. M., 1978, "On a Direct Algorithm for Nonlinear Complementarity Problems," *SIAM J. Control and Optimization*, Vol. 16, pp. 504–511.

Koppers Company, 1975, *Engineer's Handbook*, 9th ed., p. 17.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

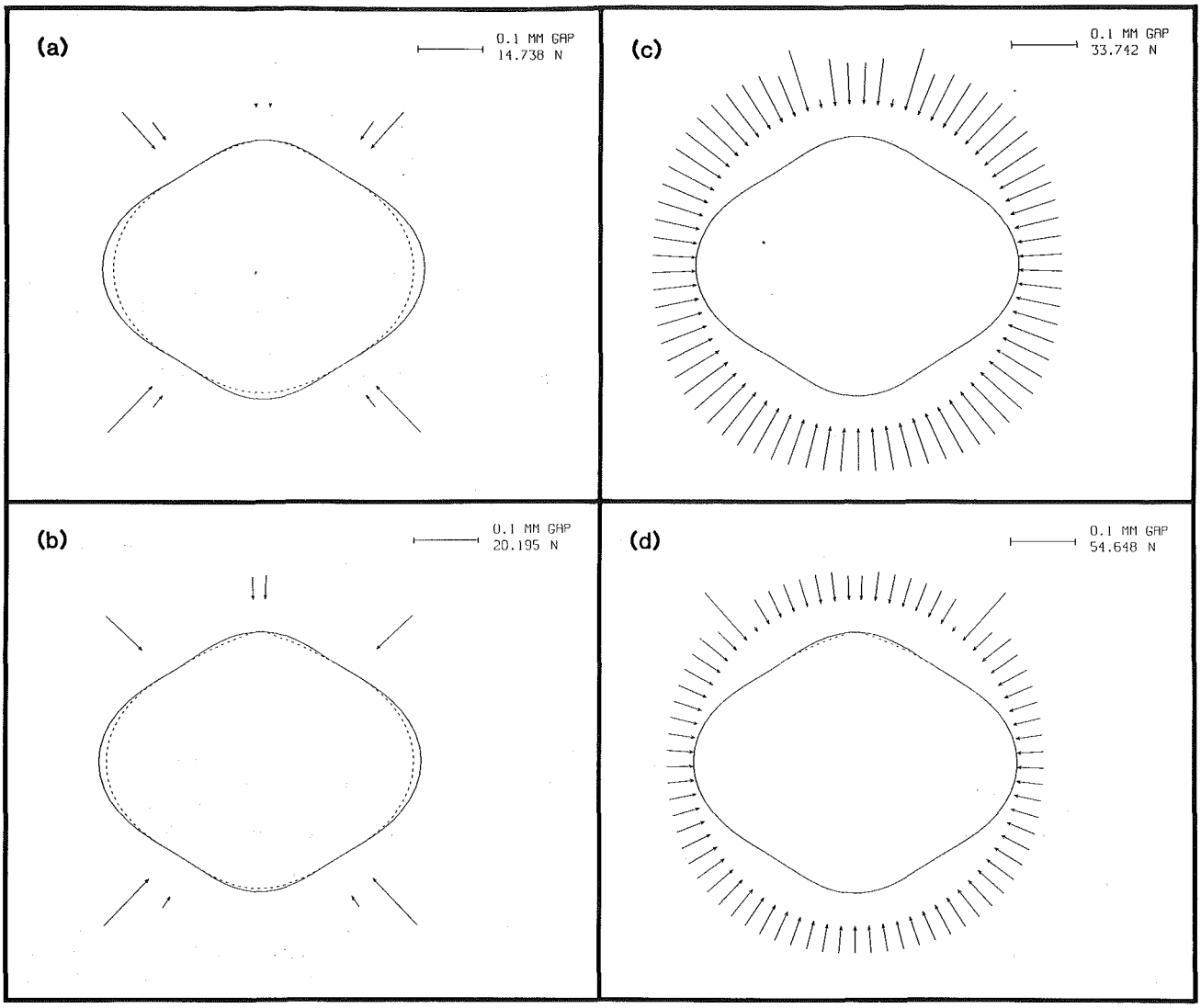


Fig. 14 Contact pattern and loads—handbook ring with distorted bore, θ_s = 90 deg

Kostreva, M. M., 1984, "Elasto-hydrodynamic Lubrication: A Non-linear Complementarity Problem," *International Journal for Numerical Methods in Fluids*, Vol. 4, pp. 377–397.

Mierbach, A., Duck, G. E., and Newman, B. A., 1983, "Heat Flow Through Piston Rings and Its Influence on Shape," SAE Paper 831283.

Murty, K. G., 1976, *Linear and Combinatorial Programming*, John Wiley and Sons, New York, Chapter 16.

Sargent, R. W. H., 1978, "An Efficient Implementation of the Lemke Algorithm and Its Extension to Deal with Upper and Lower Bounds," *Mathematical Programming Study*, Vol. 7, pp. 36–54.

Swift, H. W., 1947, "Elastic Deformation of Piston Rings," *Engineering*, Vol. 163, pp. 161–165.

Timoshenko, S., and Lessells, J. M., 1925, *Applied Elasticity*, Westinghouse Technical Night School Press, pp. 240–242.

Truesdell, C., 1961, "Exact Theory of Self-Expanding Piston Rings," *Ingenieur-Archiv*, Vol. 30 (2nd issue), pp. 77–87.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Pure Bending of Curved Beams of Thin-Walled Rectangular Box Section

R. D. Cook
Department of Engineering Mechanics,
University of Wisconsin-Madison,
Madison, WI 53706
Mem. ASME

A closed-form solution of the subject problem is presented. The analytical method resembles that used by Bleich (1933) to study curved beams of I *or* T *section. It is found that the circumferential stress may be smaller than a perpendicular stress that arises from flexing of parts of the box. Accuracy of the solution is verified by comparison with finite element analyses.*

Introduction

We consider a curved beam bent in its own plane by moment M_o. The cross-section, depicted in Fig. 1(*a*), is thin-walled and rectangular. Due to curvature about the axis of revolution, M_o produces deformations in the plane of the cross-section: In Fig. 1(*c*), compressive circumferential stress σ_θ in the inner cylindrical flange forces it outward, while tensile σ_θ in the outer flange forces it inward. These motions are reversed if M_o is reversed. The deformation reduces the flexural stiffness seen by moment M_o, leads to a σ_θ stress substantially greater than predicted by the flexure formula My/I, and creates corner bending moments M_1 and M_2 that may produce flexural stresses σ_x almost twice the maximum σ_θ stress.

A previous analysis of the subject problem was performed by Timoshenko (1923). His analysis uses a simple assumption for the deformed shape of the cross-section, and is accurate only if radii r_1 and r_2 are comparable and dimension $2c$ does not substantially exceed dimension $r_2 - r_1$.

In what follows, the box section is provided with a closed-form solution by use of methods applied with great success by Bleich (1933) to curved beams of I and T cross-section. The resulting formulas for the box beam are easy to implement on a desktop computer, and yield the following information:

(1) circumferential stresses σ_θ, in particular at corners of the cross-section where σ_θ stresses are largest;

(2) flexural stresses σ_x associated with bending moments M_1 and M_2 at corners of the cross section; and

(3) the bending stiffness seen by the applied moment M_o.

Analysis

We begin with the two radial webs, each of thickness t_w, which deform symmetrically with respect to the plane $x = 0$ (Fig. 1(*c*)). Each web is regarded as a thin annular plate, simply supported at $r = r_1$ and at $r = r_2$, where it is loaded by bending moments M_1 and M_2. The deflection w of the web can be written in the form (Timoshenko et al., 1959)

$$w = B_1 r^2 \ln \frac{r}{r_2} + B_2 \ln \frac{r}{r_2} + B_3 r^2 + B_4 \tag{1}$$

where $B_1 - B_4$ are constants. At corners of the box, $r = r_1$ and $r = r_2$, we apply the boundary conditions $w_i = 0$ and $M_r = M_i$ $(i = 1, 2)$. With E = elastic modulus and ν = Poisson's ratio,

$$M_r = D_w \left(\frac{d^2 w}{dr^2} + \frac{\nu}{r} \frac{dw}{dr} \right) \text{ where } D_w = \frac{E t_w^3}{12(1-\nu^2)}. \tag{2}$$

Thus, we obtain four equations to be solved for $B_1 - B_4$, written here in matrix format as

$$[\mathbf{p}][B_1\ B_2\ B_3\ B_4]^T = \left[0\ 0\ \frac{M_1}{D_w}\ \frac{M_2}{D_w} \right]^T. \tag{3}$$

With the notion $[\mathbf{s}] = [\mathbf{p}]^{-1}$, equations (1) and (3) yield the rotations at $r = r_1$ and $r = r_2$

$$\left(\frac{dw}{dr} \right)_1 = f_1 \frac{M_1}{D_w} + f_2 \frac{M_2}{D_w} \qquad \left(\frac{dw}{dr} \right)_2 = f_3 \frac{M_1}{D_w} + f_4 \frac{M_2}{D_w} \tag{4}$$

where

$$f_1 = r_1 \left(1 + 2 \ln \frac{r_1}{r_2} \right) s_{13} + \frac{1}{r_1} s_{23} + 2 r_1 s_{33} \tag{5a}$$

$$f_2 = r_1 \left(1 + 2 \ln \frac{r_1}{r_2} \right) s_{14} + \frac{1}{r_1} s_{24} + 2 r_1 s_{34} \tag{5b}$$

$$f_3 = r_2 s_{13} + \frac{1}{r_2} s_{23} + 2 r_2 s_{33} \tag{5c}$$

$$f_4 = r_2 s_{14} + \frac{1}{r_2} s_{24} + 2 r_2 s_{34}. \tag{5d}$$

The foregoing solution for the annular web must be matched to solutions for the cylindrical flanges.

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 9, 1989; final revision, Feb. 13, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The inner and outer flanges are assumed to behave as thin cylindrical shells. With u the radial displacement, the governing equation for each ($i = 1, 2$) is (Timoshenko et al., 1959)

$$D_i \frac{d^4u_i}{dx^4} - p = 0 \text{ where } D_i = \frac{Et_i^3}{12(1-\nu^2)}. \tag{6}$$

Here, p is an outward radial pressure associated with external loading and membrane stiffness of the shell. Let σ_θ be the circumferential membrane stress at a flange midsurface, $r = r_1$ or $r = r_2$. With $\bar{\sigma}_\theta$ the value of σ_θ at a corner, $x = \pm c$, the value of σ_θ at arbitrary x is

$$\sigma_\theta = \bar{\sigma}_\theta + E\frac{u}{r}. \tag{7}$$

We assume that radial displacement u is zero at the corners. The radial component of σ_θ is $p = -\sigma_\theta t/r$, so that equations (6) and (7) yield

$$D_i \frac{d^4u_i}{dx^4} + \frac{Et_i}{r_i^2} u_i = -\frac{\bar{\sigma}_{\theta i} t_i}{r_i} \tag{8}$$

where $i = 1, 2$. With terms antisymmetric about $x = 0$ discarded, equation (8) has the solutions

$$u_1 = -(\bar{\sigma}_{\theta 1} r_1/E) + C_1 \sin\alpha_1 x \sinh\alpha_1 x + C_2 \cos\alpha_1 x \cosh\alpha_1 x \tag{9a}$$

$$u_2 = -(\bar{\sigma}_{\theta 2} r_2/E) + C_3 \sin\alpha_2 x \sinh\alpha_2 x + C_4 \cos\alpha_2 x \cosh\alpha_2 x \tag{9b}$$

for inner and outer flanges, respectively, where $C_1 - C_4$ are constants and

$$\alpha_i^4 = \frac{3(1-\nu^2)}{r_i^2 t_i^2} \text{ where } i = 1,2. \tag{10}$$

To connect the flanges to the web, we match rotations

$$\left(\frac{du_i}{dx}\right)_{x=c} = \left(\frac{dw}{dr}\right)_{r=r_i} \text{ for } i = 1,2 \tag{11}$$

and also match bending moments

$$M_1 = D_1\left(\frac{d^2u_1}{dx^2}\right)_{x=c} \text{ and } M_2 = -D_2\left(\frac{d^2u_2}{dx^2}\right)_{x=c}. \tag{12}$$

From equations (4), (5), (9), (11), and (12) we obtain

$$[\mathbf{g}][C_1\ C_2\ C_3\ C_4]^T = \left[\frac{\bar{\sigma}_{\theta 1}}{E}\ \frac{\bar{\sigma}_{\theta 2}}{E}\ 0\ 0\right]^T. \tag{13}$$

In matrix [g], $g_{13} = g_{14} = g_{21} = g_{22} = 0$. As for the remaining g_{ij}, with the definitions

$$a_1 = \sin\alpha_1 c \sinh\alpha_1 c \qquad a_2 = \cos\alpha_1 c \cosh\alpha_1 c$$
$$a_3 = \sin\alpha_2 c \sinh\alpha_2 c \qquad a_4 = \cos\alpha_2 c \cosh\alpha_2 c$$
$$f_5 = 2\alpha_1^2 \frac{D_1}{D_w} \qquad f_6 = 2\alpha_2^2 \frac{D_2}{D_w} \qquad f_7 = \ln\frac{r_2}{r_1}$$
$$f_8 = (\sin\alpha_1 c \cosh\alpha_1 c - \cos\alpha_1 c \sinh\alpha_1 c)(t_1/r_1\alpha_1) \tag{14}$$
$$f_9 = (\sin\alpha_1 c \cosh\alpha_1 c + \cos\alpha_1 c \sinh\alpha_1 c)(t_1/r_1\alpha_1)$$
$$f_{10} = (\sin\alpha_2 c \cosh\alpha_2 c - \cos\alpha_2 c \sinh\alpha_2 c)(t_2/r_2\alpha_2)$$
$$f_{11} = (\sin\alpha_2 c \cosh\alpha_2 c + \cos\alpha_2 c \sinh\alpha_2 c)(t_2/r_2\alpha_2)$$

the rest of the g_{ij} are

$$g_{11} = \frac{a_1}{r_1} \qquad g_{12} = \frac{a_2}{r_1} \qquad g_{23} = \frac{a_3}{r_2} \qquad g_{24} = \frac{a_4}{r_2}$$
$$g_{31} = \frac{r_1\alpha_1^2}{t_1} f_9 - a_2 f_1 f_5 \qquad g_{32} = -\frac{r_1\alpha_1^2}{t_1} f_8 + a_1 f_1 f_5 \tag{15}$$
$$g_{33} = a_4 f_2 f_6 \qquad g_{34} = -a_3 f_2 f_6 \qquad g_{41} = -a_2 f_3 f_5 \qquad g_{42} = a_1 f_3 f_5$$
$$g_{43} = \frac{r_2\alpha_2^2}{t_2} f_{11} + a_4 f_4 f_6 \qquad g_{44} = -\frac{r_2\alpha_2^2}{t_2} f_{10} - a_3 f_4 f_6.$$

Solution of equations (13) yields constants $C_1 - C_4$, hence stress σ_θ in a flange is obtained from equations (7) and (9). Forces F_1 and F_2, carried by the width $2c$ of the respective flanges, are

$$F_1 = 2\int_0^c \sigma_{\theta 1} t_1 dx = f_{12}\bar{\sigma}_{\theta 1} + f_{14}\bar{\sigma}_{\theta 2} \tag{16a}$$

$$F_2 = 2\int_0^c \sigma_{\theta 2} t_2 dx = f_{13}\bar{\sigma}_{\theta 1} + f_{15}\bar{\sigma}_{\theta 2} \tag{16b}$$

where

$$f_{12} = f_8 q_{11} + f_9 q_{21} \qquad f_{13} = f_{10} q_{31} + f_{11} q_{41}$$
$$f_{14} = f_8 q_{12} + f_9 q_{22} \qquad f_{15} = f_{10} q_{32} + f_{11} q_{42} \tag{17}$$

and the q_{ij} are coefficients of the inverse matrix $[\mathbf{q}] = [\mathbf{g}]^{-1}$.

As for σ_θ in the radial webs, we assume the distribution used in the analysis of solid curved beams in bending, that is (Cook, et al., 1985) $\sigma_{\theta w} = C_5 + C_6/r$. To evaluate constants C_5 and C_6, we require that $\sigma_{\theta w} = \bar{\sigma}_{\theta 1}$ at $r = r_1$ and $\sigma_{\theta w} = \bar{\sigma}_{\theta 2}$ at $r = r_2$. Thus,

$$\sigma_{\theta w} = \frac{\bar{\sigma}_{\theta 2} r_2 - \bar{\sigma}_{\theta 1} r_1}{r_2 - r_1} + \frac{r_1 r_2 (\bar{\sigma}_{\theta 1} - \bar{\sigma}_{\theta 2})}{r(r_2 - r_1)}. \tag{18}$$

Because the total circumferential force must vanish,

$$F_1 + F_2 + 2\int_{r_1}^{r_2} \sigma_{\theta w} t_w dr = 0. \tag{19}$$

Equations (16)–(19) yield

$$f_{16}\bar{\sigma}_{\theta 1} + f_{17}\bar{\sigma}_{\theta 2} = 0 \tag{20}$$

where, with $a_5 = 2t_w$ and $a_6 = a_5 r_1 r_2 f_7/(r_2 - r_1)$,

$$f_{16} = f_{12} + f_{13} - a_5 + a_6 \qquad f_{17} = f_{14} + f_{15} + a_5 - a_6. \tag{21}$$

The total moment provided by stresses σ_θ in flanges and webs must equal the applied moment M_o. Because of equation (19), the axis about which moments are taken is immaterial. Accordingly, choosing the axis of revolution,

$$F_1 r_1 + F_2 r_2 + 2\int_{r_1}^{r_2} \sigma_{\theta w} t_w r\, dr = M_o. \tag{22}$$

Equations (16), (17), (18), and (22) yield

$$\left(f_{18} - \frac{f_{16}}{f_{17}} f_{19}\right)\bar{\sigma}_{\theta 1} = M_o \tag{23}$$

where, with $a_7 = t_w(r_1 + r_2)$ and $a_8 = 2t_w r_1 r_2$,

$$f_{18} = r_1 f_{12} + r_2 f_{13} - a_7 r_1 + a_8 \qquad f_{19} = r_1 f_{14} + r_2 f_{15} + a_7 r_2 - a_8. \tag{24}$$

Having determined $\bar{\sigma}_{\theta 1}$ at inner corners from equation (23), equation (20) yields $\bar{\sigma}_{\theta 2}$ at outer corners. Bending moments, from equations (9), (12), (13), and (14) are found to be

$$M_1 = 2D_1\alpha_1^2(C_1 a_2 - C_2 a_1) \qquad M_2 = 2D_2\alpha_2^2(C_4 a_3 - C_3 a_4) \tag{25}$$

where $C_1 - C_4$ are determined from equation (13). Axially directed flexural stresses are determined by the flexure formula; for example, in the flange at an inner corner, $\sigma_x = 6M_1/t_1^2$.

Because we have assumed that radial displacement u is zero at the corners, the relative rotation between radial faces separated by an angle θ_o, Fig. 1(*b*), is

$$\phi_B = \frac{(\theta_o r_2)\bar{\epsilon}_{\theta 2} - (\theta_o r_1)\bar{\epsilon}_{\theta 1}}{r_2 - r_1} = \frac{r_2\bar{\sigma}_{\theta 2} - r_1\bar{\sigma}_{\theta 1}}{E(r_2 - r_1)}\theta_o. \tag{26}$$

If the box beam were straight rather than curved, then a cross-section would not deform in its own plane, and the rotation could be easily determined from the flexure formula as

$$\phi_S = \frac{M_o(\theta_o r_c)}{EI_c} \tag{27}$$

where $\theta_o r_c$ is regarded as the centroidal length of the beam and I_c is the centroidal moment of inertia of its cross-section.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Results for circumferential stress and cross-section rotation. All results are relative to analogous straight-beam values for the corresponding cross-section.

Section	Present Analysis $\bar{\sigma}_{\theta 1}$	Present Analysis ϕ_B	Finite Elements $\bar{\sigma}_{\theta 1}$	Finite Elements ϕ_B	Timoshenko (1923) $\bar{\sigma}_{\theta 1}$	Timoshenko (1923) ϕ_B
Fig. 2(*a*)	1.68	1.62	1.64	1.59	1.58	1.58
Fig. 2(*b*)	4.53	2.79	4.21	3.07	2.40	2.40
Fig. 2(*c*)	4.29	4.15	4.10	3.96	2.23	2.23
Fig. 2(*d*)	7.91	4.84	7.37	5.19	2.23	2.23

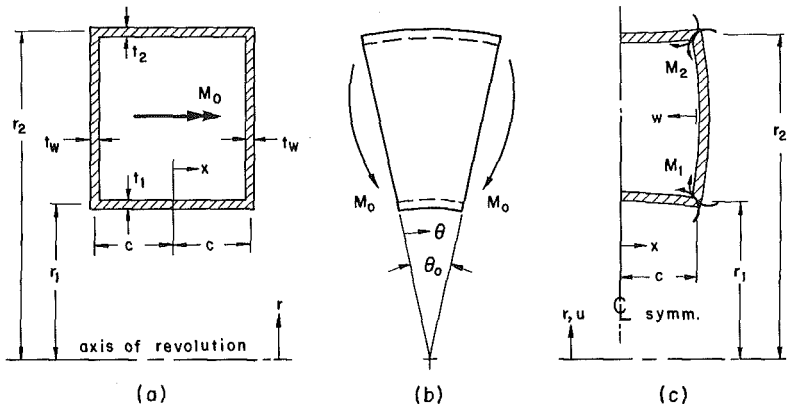


Fig. 1 (*a*) Cross-section of thin-walled curved box beam. Radii r_1 and r_2 locate midsurfaces of the cylindrical flanges. (*b*) View parallel to the axis of revolution. (*c*) Deformation of the cross-section, showing bending moments M_1 and M_2 that arise at the corners.

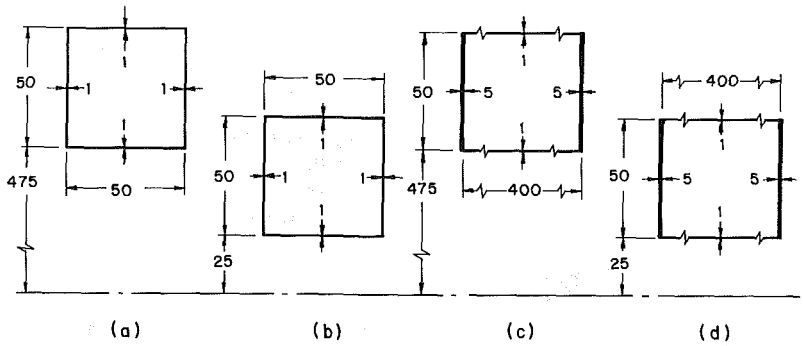


Fig. 2 Cross-sections used in numerical tests

Results

For illustration and verification, the foregoing analysis is applied to the four cross-sections depicted in Fig. 2. For comparison, the same sections are analyzed by the finite element method discussed by Cook (1989) and the approximate method of Timoshenko (1923). Admittedly, the latter method is not intended for use unless $r_2 - r_1$ is appreciably greater than r_1.

Results for circumferential stress $\bar{\sigma}_{\theta 1}$ and cross-sectional rotation ϕ_B are reported in Table 1. All results are reported as *ratios*. Thus, all stresses $\bar{\sigma}_{\theta 1}$ have been divided by the stress given by the straight-beam flexure formula, $\sigma = M_o(r_2 - r_1)/2I_c$ for the cross-section at hand. Similarly, values of rotation ϕ_B have been divided by ϕ_S from equation (27).

We see that straight-beam theory seriously underestimates both stresses and cross-section rotations, as does the Timoshenko formula to a lesser degree. These underestimates arise because straight-beam theory ignores deformations of the cross-section, Timoshenko theory misrepresents the deformations if the cross-section is wide or sharply curved, and both theories assume that σ_θ varies linearly with r rather than as described by equation (18). In comparison with finite element results, the present theory may overestimate $\bar{\sigma}_{\theta 1}$ by almost 10 percent, and may underestimate rotation (i.e., overestimate stiffness) by almost 10 percent. We think that the discrepancy arises mainly because the present theory imposes $u = 0$ at $x = \pm c$ and does not account for radial stress σ_r in the webs.

As for flexural stresses σ_x associated with corner bending moment M_1, according to the present analysis the σ_x stresses at inner corners are the following percentages of $\bar{\sigma}_{\theta 1}$ in the respective parts of Fig. 2: 50 percent in Fig. 2(*a*), 17 percent in Fig. 2(*b*), 186 percent in Fig. 2(*c*), and 182 percent in Fig. 2(*d*). From finite element analysis, the corresponding percentages are 49 percent, 14 percent, 175 percent, and 164 percent. Clearly, these flexural stresses can be the largest stress in the beam if the radial webs are substantially thicker than the flanges.

References

Bleich, H., 1933, "Die Spannungsverteilung in den gurtungen gekrümmter Stabe mit T- und I-förmigem Querschnitt," *Der Stahlbau* (appendix to *Die Bautechnik*), Vol. 6, No. 1, pp. 3–6.

Cook, R. D., 1989, "Axisymmetric Finite Element Analysis for Pure Moment Loading of Curved Beams and Pipe Bends," *Computers and Structures*, Vol. 33, No. 2, pp. 483–487.

Cook, R. D., and Young,W. C., 1985, *Advanced Mechanics of Materials*, Macmillan, New York, pp. 411–414.

Timoshenko, S., 1923, "Bending Stresses in Curved Tubes of Rectangular Cross Section," *Trans. ASME*, Vol. 45, pp. 135–140.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, pp. 51–54, 466–476.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

E. Madenci
Department of Aerospace and Mechanical Engineering,
University of Arizona,
Tuscon, AZ 85721
Mem. ASME

R. A. Westmann
Mechanical, Aerospace and Nuclear Engineering Department,
University of California,
Los Angeles, CA 90024-1597
Mem. ASME

Local Delamination Buckling in Layered Systems

This paper presents an analytical solution to the problem of local buckling induced by delamination of a layered plate. Delamination growth and buckling is an observed failure mode in laminated structures subjected to compressive loads. Previous analytical studies of the phenomenon rest upon simplifying structural and geometric approximations. The purpose of this paper is to present solutions for this problem using the classical three-dimensional theory of elasticity to predict the buckled equilibrium state. Solutions to the problem of a plate with a circular delamination subjected to axisymmetric and uniaxial in-plane compressive loadings are obtained using mathematical techniques appropriate for mixed boundary value problems.

I Introduction

Laminated plates are formed by adhesively bonding multiple thin sheets of homogeneous isotropic metals. The intent is that the resulting laminated structural component will exhibit greater damage tolerance due to the high value of fracture toughness possessed by each thin lamina.

Such layered systems are susceptible to interface delamination arising from either manufacturing errors resulting in poor bonding or from low-velocity lateral impact. While such damage may not seriously degrade the tensile strength properties of the laminate, concern about the compressive strength of the partially delaminated structure is most valid. Should the delamination be sufficiently large, local buckling in the neighborhood of the damage will lead to delamination growth and/or catastrophic failure (Chai, Knauss, and Babcock, 1983; McCarty and Ratwani, 1983; Starnes and Williams, 1984).

Because of the importance of this subject, several studies have been undertaken to predict the level of compressive load resulting in plate buckling induced by delaminations (Kachanov, 1976; Chai, Babcock, and Knauss, 1981; Chai and Babcock, 1985; Bottega and Maewal, 1983; Simitses, Sallam, and Yin, 1984; Yin, Sallam, and Simitses 1984; Yin and Fei 1985; Yin and Wang 1984). To date, these investigations have utilized structural mechanics theories such as those for beams and plates and, therefore, employed approximations with respect to the boundary conditions enforced at the edges of the delamination.

Such approximate solutions are of interest as they supply estimates of the critical delamination size as well as other practical information. However, solutions using a more exact theory are preferable because they provide more accurate results and give a basis for assessing the accuracy of simpler, more tractable models.

This paper presents the solutions to two specific boundary value problems using a three-dimensional elasticity model not previously applied to this class of problems. Employing the stability equations of linear elasticity theory to predict the buckling response of the stated problem leads to a standard mixed boundary value problem. An example of this approach by Keer, Nemat-Nasser, and Oranratnachai (1982) estimates the buckling load of an elastic half-plane and a layer, each containing an array of coplanar, equally spaced cracks. The results presented by Keer et al. (1982) are for a class of hypoelastic materials for which the Jaumann rate of the Kirchoff stress is related to the deformation rate tensor by an isotropic linear relation similar to Hooke's law.

Section II presents the mathematical model for the two-dimensional loading conditions of axisymmetric and uniaxial in-plane compressive loading. In Section III the solution for the case of axisymmetric loading is obtained by first reducing the problem to a pair of dual integral equations and then to the solution of a single homogeneous Fredholm integral equation of the second kind.

Section IV is concerned with the more difficult case of uniaxial loading. A mathematical approximation permits the problem to be stated in terms of two pairs of simultaneous dual integral equations. These in turn are converted to the solution of a pair of coupled, homogeneous Fredholm integral equations of the second kind.

Section V outlines the numerical method used to solve these two eigenvalue problems. Numerical results are presented to show how the buckling stress depends upon the applicable geometric parameters.

II Problem Statement

This paper is concerned with the stability of a flat elastic layer containing a circular delamination and subjected to in-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, Dallas, Texas, Nov. 25–30, 1990.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 22, 1988; final revision, Jan. 15, 1990.

Paper No. 90-WA/APM-40.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

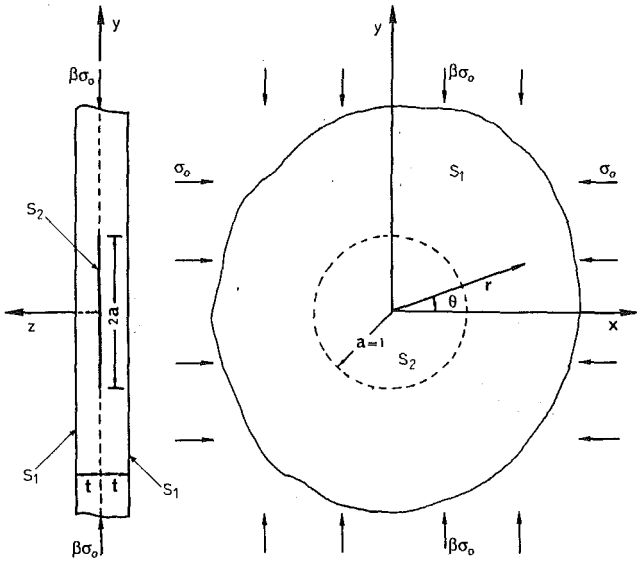


Fig. 1 Layered structure with circular delamination

plane compressive loads. A cylindrical coordinate system (r, θ, z) is employed in which the plane $z = 0$ coincides with the midsurface of the layer, as shown in Fig. 1. The total layer thickness is taken to be $2t$, and the bounding surface S_1 of the layer becomes

$$S_1 = \{r, \theta, z: r\epsilon[0,\infty),\ \theta\epsilon[0,2\pi],\ z = \pm t\}.$$

The circular delamination, radius $a = 1$, is located in the midplane of the layer and centered at $r = 0$; the delamination surface S_2 is therefore described by

$$S_2 = \{r, \theta, z: r\epsilon[0, a),\ \theta\epsilon[0, 2\pi],\ z = \pm 0\}.$$

The layer is composed of a homogeneous, elastic, and isotropic material with shear modulus, μ, and Poisson's ratio, ν.

The plate is subjected to an in-plane compressive loading σ_o resulting in the initial equilibrium stress state with surfaces S_1 and S_2 remaining traction-free. The problem then posed is to determine the value of σ_o for which the delaminated layer buckles. Solutions are sought for the buckled equilibrium state (adjacent to the initial state) and are to be obtained within the scope of the three-dimensional equations of elastic stability (Flugge, 1972; Washizu, 1975; Bolotin, 1963).

The solution to this problem provides a good approximation for the practical case of a laminated plate consisting of a finite number of homogeneous and elastic layers with perfect bonding, except for a circular interface delamination.

III Solution for Axisymmetric Compression

Under axisymmetric compression, $\beta = 1$ in Fig. 1, the displacement equilibrium equations of elastic stability (Flugge, 1972) may be written as

$$\frac{2(1-\nu)}{1-2\nu}\left[\frac{1}{r}(ru_r)_{,r}+u_{z,z}\right]_{,r}-(u_{z,r}-u_{r,z})_{,z}-\frac{-\sigma_o}{\mu}\left[\frac{1}{r}(ru_r)_{,r}\right]_{,r}=0 \tag{1}$$

$$\frac{2(1-\nu)}{1-2\nu}\left[\frac{1}{r}(ru_r)_{,r}+u_{z,z}\right]_{,z}+\frac{1}{r}[r(u_{z,r}-u_{r,z})]_{,r}-\frac{-\sigma_o}{\mu}\frac{1}{r}(ru_{z,r})_{,r}=0$$

where u_r, u_θ, and u_z are the components of the displacement field referenced to the cylindrical coordinates (r, θ, z). Due to the anticipated axial symmetry of the buckled state corresponding to the lowest value of σ_o, the displacement component u_θ and derivatives, with respect to θ, vanish in equation (1).

Boundary conditions associated with the traction-free surface S_1 then become

$$\left.\begin{aligned}&\frac{2\mu}{(1-2\nu)}\left[(1-\nu)u_{z,z}+\nu\frac{1}{r}(ru_r)_{,r}\right]=0 &(2a)\\ &\mu[u_{r,z}+u_{z,r}]\doteq 0 &(2b)\end{aligned}\right\} z=\pm t,\ r\epsilon[0,\infty).$$

Symmetry of the solution with respect to the midplane ($z = 0$) and the requirement that the surface S_2 (delamination) be traction-free result in the following boundary conditions:

$$\mu[u_{r,z}+u_{z,r}]=0 \qquad z=0+,\ r\epsilon[0,\infty) \tag{3a}$$

$$\frac{2\mu}{(1-2\nu)}\left[(1-\nu)u_{z,z}+\nu\frac{1}{r}(ru_r)_{,r}\right]=0 \qquad z=0+,\ r\epsilon[0,1) \tag{3b}$$

$$u_z=0 \qquad z=0+,\ r\epsilon(1,\infty). \tag{3c}$$

Further, the far-field regularity conditions require that

$$u_r,\ u_z\to 0 \text{ for } z\epsilon[0,t],\ r\to\infty.$$

The solution procedure begins by representing the displacement field in terms of the following integrals[1] (Harding and Sneddon, 1945)

$$u_r(r,z)=\int_0^\infty \Psi_2(z,\xi)J_1(r\xi)d\xi$$
$$u_z(r,z)=\int_0^\infty \Psi_1(z,\xi)J_0(r\xi)d\xi \tag{4}$$

where $\Psi_1(z, \xi)$, $\Psi_2(z, \xi)$ are unknown auxiliary functions. Substitution from equation (4) into equation (1) reduces the partial differential equations (1) to a system of two, coupled, second-order ordinary differential equations for $\Psi_1(z, \xi)$ and $\Psi_2(z, \xi)$[2].

$$\Psi_2''-\frac{2(1-\nu)}{1-2\nu}\xi^2\Psi_2-\frac{1}{1-2\nu}\xi\Psi_1'+\frac{\sigma_o}{\mu}\xi^2\Psi_2=0$$
$$\frac{2(1-\nu)}{1-2\nu}\Psi_1''-\xi^2\Psi_1+\frac{1}{1-2\nu}\xi\Psi_2'+\frac{\sigma}{\mu}o\xi^2\Psi_1=0 \tag{5}$$

Nontrivial solutions to equations (5) are readily determined to be

$$\Psi_1(z,\xi)=C_1\,\mathrm{Cosh}[\tau_1\xi(t-z)]-C_2\,\mathrm{Sinh}[\tau_1\xi(t-z)]+C_3\,\mathrm{Cosh}[\tau_2\xi(t-z)]-C_4\,\mathrm{Sinh}[\tau_2\xi(t-z)] \tag{6}$$
$$\Psi_2(z,\xi)=-\tau_1^{-1}\{C_2\,\mathrm{Cosh}[\tau_1\xi(t-z)]-C_1\,\mathrm{Sinh}[\tau_1\xi(t-z)]\}-\tau_2\{C_4\,\mathrm{Cosh}[\tau_2\xi(t-z)]-C_3\,\mathrm{Sinh}[\tau_2\xi(t-z)]\}$$

where $\tau_1^2=1-\frac{1-2\nu}{2(1-\nu)}\frac{\sigma_o}{\mu}$, $\tau_2^2=1-\frac{\sigma_o}{\mu}$. The four coefficients $C_1 - C_4$ remain to be determined from the boundary conditions (2) and (3).

Enforcement of the traction-free condition on the surface $z = t$ equations (2a), (2b) and the shear-free condition in the midplane $z = 0$ equation (3a) permits the expression of C_1, C_2, C_3 in terms of C_4. The remaining unknown, C_4, is determined by imposing the mixed boundary conditions equations (3b), (3c) in the plane of the delamination region.

[1] In this paper $J_\nu(x)$ denotes the Bessel function of the first kind, argument x and order ν.

[2] Prime denotes differentiation with respect to variable z throughout this paper.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Setting

$$C(\xi)=\frac{\tau\ \mathrm{Sinh}(\tau_1 t\xi)\ \mathrm{Cosh}(\tau_2 t\xi)-\mathrm{Cosh}(\tau_1 t\xi)\ \mathrm{Sinh}(\tau_2 t\xi)}{\mathrm{Cosh}(\tau_1 t\xi)-\mathrm{Cosh}(\tau_2 t\xi)}\,C_4$$

and enforcing equations (3*b*), (3*c*) result in the following pair of dual integral equations

$$\int_0^\infty C(\xi)\,\xi\,J_o(r\xi)d\xi=\int_0^\infty G\left(\xi;\frac{\sigma_o}{\mu},t\right)C(\xi)\,\xi\,J_o(r\xi)d\xi\quad r\epsilon[0,1) \tag{7a}$$

$$\int_0^\infty C(\xi)\,J_o(r\xi)d\xi=0\qquad r\epsilon[1,\infty) \tag{7b}$$

where the known function $G(\xi;\frac{\sigma_o}{\mu},t)$ is

$$G\left(\xi;\frac{\sigma_o}{\mu},t\right)=\frac{1}{1-\tau}\times\frac{[4\tau+(1-\tau)^2e^{-(\tau_1+\tau_2)t\xi}-(1+\tau)e^{-(\tau_1-\tau_2)t\xi}-\tau(1+\tau)e^{(\tau_1\tau_2)t\xi}}{(1+\tau)\ \mathrm{Sinh}[(\tau_1+\tau_2)t\xi]-(1-\tau)\ \mathrm{Sinh}[(\tau_1+\tau_2)t\xi]}$$

and

$$\tau=4\tau_1\tau_2/(1+\tau_2^2)^2.$$

The dual integral equations (7) are converted to a homogeneous Fredholm integral equation of the second kind (Sneddon and Lowengrub, 1969) by means of the following integral representation for $C(\xi)$ (Copson, 1961):

$$C(\xi)=\xi^{1/2}\int_0^1\alpha^{1/2}\Psi(\alpha)\,J_{1/2}(\xi\alpha)\,d\alpha. \tag{8}$$

In this representation, $\Psi(\alpha)$ is assumed to be continuous over the interval [0, 1] and is required to satisfy the condition

$$\lim_{a\to 0+}[\alpha^{1/2}\psi(\alpha)]=0.$$

This representation of $C(\xi)$ identically satisfies equation (7*b*). Integrating equation (8) by parts and substituting for $C(\xi)$ into equation (7*a*) lead to the following Abel integral equation:

$$\int_0^r\frac{d[\Psi(\alpha)]}{d\alpha}\frac{d\alpha}{\sqrt{r^2-\alpha^2}}=\sqrt{\frac{\pi}{2}}\int_0^\infty G\left(\xi;\frac{\sigma_o}{\mu},t\right)C(\xi)\xi J_o(r\xi)d\xi. \tag{9}$$

Applying Abel's solution method (Whittaker and Watson, 1920) to invert the left-hand side of equation (9), and then substituting from equation (8), give the final governing integral equation for $\Psi(\alpha)$

$$\Psi(\alpha)=\lambda^*\int_0^1 K\left(\alpha,\eta;\frac{\sigma_o}{\mu},t\right)\Psi(\eta)\ d\eta\quad(\lambda^*=1). \tag{10}$$

In equation (10) the symmetric kernel of the homogeneous Fredholm integral equation of the second kind is specifically written as

$$K\left(\alpha,\eta;\frac{\sigma_o}{\mu},t\right)=\frac{2}{\pi}\int_0^\infty G\left(\xi;\frac{\sigma_o}{\mu},t\right)\sin(\alpha\xi)\sin(\eta\xi)d\xi. \tag{11}$$

Nontrivial solutions to equation (10) are obtained by searching for those values of $\frac{\sigma_o}{\mu}$ such that $\lambda^*=1$ is an eigenvalue.

IV Solution for Uniaxial Compression

In the case of uniaxial compression ($\beta=0$ in Fig. 1), the governing equations for the buckled state in cylindrical coordinates become

$$\nabla^2u_r+\frac{1}{(1-2\nu)}\Delta_{,r}-\frac{1}{r^2}(2u_{\theta,\theta}+u_r)-\frac{\sigma_o}{2\mu}\left[\left(\frac{1}{r}(ru_r)_{,r}\right)_{,r}+\frac{1}{r^2}\left(u_{r,\theta}-2u_\theta\right)_{,\theta}\right]-\frac{\sigma_o}{2\mu}\left[\left(r\left(\frac{1}{r}u_r\right)_{,r}\right)_{,r}-\frac{1}{r^2}\left(u_{r,\theta}-2u_\theta\right)_{,\theta}\right]\mathrm{Cos}(2\theta)+\frac{\sigma_o}{2\mu}\left[2\left(\frac{1}{r}(u_{r,\theta}-u_\theta)\right)_{,r}\right]\mathrm{Sin}(2\theta)=0 \tag{12a}$$

$$\nabla^2u_\theta+\frac{1}{(1-2\nu)}\frac{1}{r}\Delta_{,\theta}+\frac{1}{r^2}(2u_{r,\theta}+u_\theta)-\frac{\sigma_o}{2\mu}\left[\left(\frac{1}{r}(ru_\theta)_{,r}\right)_{,r}+\frac{1}{r^2}(u_{\theta,\theta}+2u_r)_{,\theta}\right]-\frac{\sigma_o}{2\mu}\left[\left(r\left(\frac{1}{r}u_\theta\right)_{,r}\right)_{,r}-\frac{1}{r^2}(u_{\theta,\theta}+2u_r)_{,\theta}\right]\mathrm{Cos}(2\theta)+\frac{\sigma_o}{2\mu}\left[2\frac{1}{r}(u_{\theta,\theta}+u_r)_{,r}\right]\mathrm{Sin}(2\theta)=0 \tag{12b}$$

$$\nabla^2u_z+\frac{1}{(1-2\nu)}\Delta_{,z}-\frac{\sigma_o}{2\mu}\left[\frac{1}{r}(ru_{z,r})_{,r}+\frac{1}{r^2}\mu_{z,\theta\theta}\right]-\frac{\sigma_o}{2\mu}\times\left[r\left(\frac{1}{r}u_{z,r}\right)_{,r}-\frac{1}{r^2}u_{z,\theta\theta}\right]\mathrm{Cos}2\theta+\frac{\sigma_o}{2\mu}\left[2\left(\frac{1}{r}u_{z,\theta}\right)_{,r}\right]\mathrm{Sin}(2\theta)=0 \tag{12c}$$

in which

$$\Delta=u_{r,r}+\frac{1}{r}u_r+\frac{1}{r^2}u_{\theta,\theta}+u_{z,z}$$

$$\nabla^2=\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\partial^2}{\partial z^2}.$$

The traction-free boundary conditions on S_1 are stated as follows:

$$\left.\begin{aligned}&\frac{2\mu}{(1-2\nu)}\left\{(1-\nu)u_{z,z}+\nu\left[u_{r,r}+\frac{1}{r}(u_{\theta,\theta}+u_r)\right]\right\}=0 &(13a)\\ &\mu[u_{r,z}+u_{z,r}]=0 &(13b)\\ &\mu\left[u_{\theta,z}+\frac{1}{r}u_{z,\theta}\right]=0 &(13c)\end{aligned}\right\}\quad z=\pm t,\quad r\epsilon[0,\infty).$$

The remaining boundary conditions are imposed in order to satisfy the requirements of traction-free surfaces on S_2 and assumed symmetry conditions with respect to the midplane ($z=0$). Thus,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\left.\begin{aligned}
&\mu[u_{r,z}+u_{z,r}]=0 && r\epsilon[0,\infty) && (13d)\\
&\mu\left[u_{\theta,z}+\frac{1}{r}\,u_{z,\theta}\right]=0 && r\epsilon[0,\infty) && (13e)\\
&\frac{2\mu}{(1-2\nu)}\left\{(1-\nu)u_{z,z}+\nu\left[u_{r,r}+\frac{1}{r}\,(u_{\theta,\theta}+u_r)\right]\right\};=0 && r\epsilon[0,1) && (13f)\\
&u_z=0 && r\epsilon[1,\infty). && (13g)
\end{aligned}\right\} z=0+$$

Finally, the far-field regularity conditions require that

$$u_r,\ u_\theta,\ u_z\to 0 \text{ for } |z|\epsilon[0,\ t],\ r\to\infty.$$

As with the axisymmetric problem in Section III, these homogeneous boundary conditions, in conjunction with the homogeneous equilibrium equations, pose an eigenvalue problem by which the buckling stress, σ_0, can be determined.

An integral-series representation (Muki, 1960) for the displacement field in equation (12) is assumed to have the form

$$u_r(r,\theta,z)=\frac{1}{2}\sum_{m=0}^{\infty}\left[\int_0^{\infty}\{[\phi_m(z)+\psi_m(z)]J_{m+1}(r\xi)\right.$$
$$\left.-[\phi_m(z)-\psi_m(z)]J_{m-1}(r\xi)\}d\xi\right]\text{Cos}\,(m\theta)$$
$$u_\theta(r,\theta,z)=\frac{1}{2}\sum_{m=0}^{\infty}\left[\int_0^{\infty}\{[\phi_m(z)+\psi_m(z)]J_{m+1}(r\xi)\right.$$
$$\left.+[\phi_m(z)-\psi_m(z)]J_{m-1}(r\xi)\}d\xi\right]\text{Sin}\,(m\theta)$$
$$u_z(r,\theta,z)=\sum_{m=0}^{\infty}\left[\int_0^{\infty}\chi_m(z)J_m(r\xi)d\xi\right]\text{Cos}\,(m\theta) \qquad (14)$$

in which $\phi_m(z,\ \xi)$, $\psi_m(z,\ \xi)$, and $\chi_m(z,\ \xi)$ are unknown auxiliary functions.

Substitution from equation (14) into equation (12) reduces the partial differential equations to an infinite system of coupled, second-order, ordinary differential equations. Examination of this system reveals that the odd and even terms of the auxiliary functions are coupled only among themselves, thereby leading to two independent systems of equations. In light of this decoupling, the system of equations (15) containing only the even terms of the auxiliary functions is retained. The other system with only the odd terms of the auxiliary functions is disregarded because its solution would result in interpenetration of the delamination surfaces which is physically unacceptable. Thus, for the even terms, the governing equations are

$$\left.\begin{aligned}
&\left[\frac{2(1-\nu)}{1-2\nu}-\frac{\sigma_o}{2\mu}\right]\phi_o\xi^2-\phi_o''+\frac{1}{1-2\nu}\,\chi_0'\xi+\frac{\sigma_o}{4\mu}\phi_2\xi^2=0\\
&\left(1-\frac{\sigma_o}{2\mu}\right)\chi_0\xi^2-\frac{2(1-\nu)}{1-2\nu}\chi_o''-\frac{1}{1-2\nu}\,\phi_o'\xi+\frac{\sigma_o}{4\mu}\,\chi_2\xi^2=0
\end{aligned}\right\} m=0$$

$$\left.\begin{aligned}
&\left[\frac{2(1-\nu)}{1-2\nu}-\frac{\sigma_o}{2\mu}\right]\phi_2\xi^2-\phi_2''+\frac{1}{1-2\nu}\,\chi_2'\xi+\frac{\sigma_o}{2\mu}\phi_o\xi^2=0\\
&\left(1-\frac{\sigma_o}{2\mu}\right)\chi_2\xi^2-\frac{2(1-\nu)}{1-2\nu}\chi_o''-\frac{1}{1-2\nu}\,\phi_o'\xi+\frac{\sigma_o}{4\mu}\chi_2\xi^2=0\\
&\left(1-\frac{\sigma_o}{2\mu}\right)\psi_2\xi^2-\psi_2''=0
\end{aligned}\right\} m=2$$

$$\left.\begin{aligned}
&\left[\frac{2(1-\nu)}{1-2\nu}-\frac{\sigma_o}{2\mu}\right]\phi_m\xi^2-\phi_m''+\frac{1}{1-2\nu}\,\chi_m'\xi+\frac{\sigma_o}{4\mu}\phi_{m-2}\xi^2+\frac{\sigma_o}{4\mu}\phi_{m+2}\xi^2=0\\
&\left(1-\frac{\sigma_o}{2\mu}\right)\chi_m\xi^2-\frac{2(1-\nu)}{1-2\nu}\,\chi_m''-\frac{1}{1-2\nu}\phi_m'\xi+\frac{\sigma_o}{4\mu}\chi_{m-2}\xi^2+\frac{\sigma_o}{4\mu}\,\chi_{m+2}\xi^2=0\\
&\left(1-\frac{\sigma_o}{2\mu}\right)\psi_m\xi^2-\psi_m''+\frac{\sigma_o}{4\mu}\psi_{m-2}\,\xi^2+\frac{\sigma_o}{4\mu}\psi_{m+2}\xi^2=0
\end{aligned}\right\} m\geq 4. \qquad (15)$$

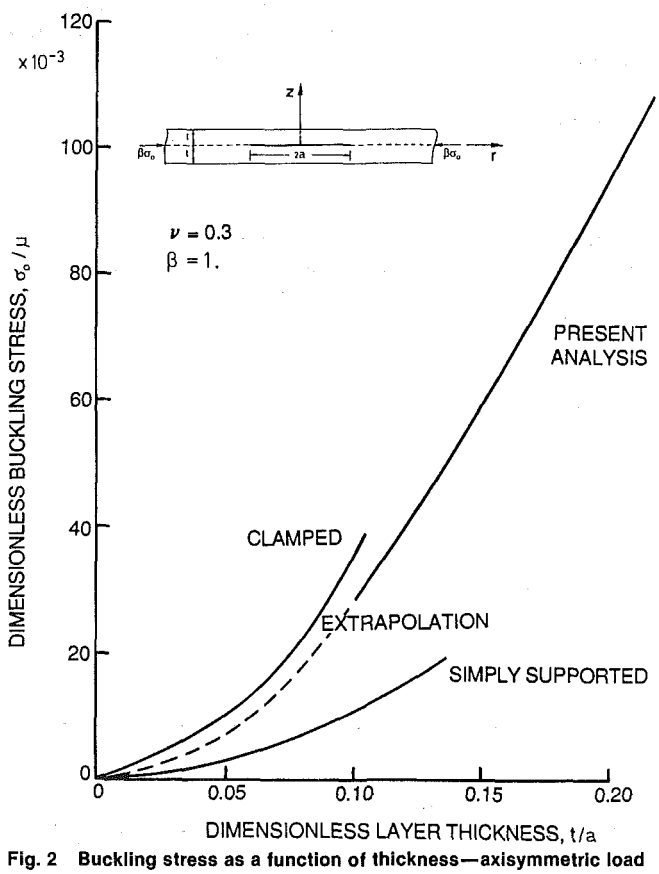


Fig. 2 Buckling stress as a function of thickness—axisymmetric load

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

To proceed the following solution form for the auxiliary functions

$$(\phi_m, \psi_m, \chi_m) = (A_m, B_m, C_m)e^{q\xi z}$$

is adopted, thereby reducing equations (15) to two uncoupled infinite systems of homogeneous algebraic equations (16) and (17).

$$\left[\begin{array}{cccc|cc|cc|} (1+\gamma) & q & \epsilon & 0 & & & & \\ q & (1-\omega) & 0 & -\epsilon & & & & \\ 2\epsilon & 0 & (1+\gamma) & q & \epsilon & 0 & & \\ 0 & -2\epsilon & q & (1-\omega) & 0 & -\epsilon & & \\ \hline & & \epsilon & 0 & (1+\gamma) & q & \epsilon & \\ & & 0 & -\epsilon & q & (1-\omega) & 0 & -\epsilon \\ \hline & & & & & & & \\ & & & & & & & \end{array}\right] \text{repeats} \left\{\begin{array}{c} A_0 \\ C_0 \\ A_2 \\ C_2 \\ A_4 \\ C_4 \\ \bullet \\ \bullet \\ \bullet \\ \bullet \end{array}\right\} = \{0\} \qquad (16)$$

$$\left[\begin{array}{c|cc|} \gamma & 0 & \\ \hline \epsilon & \gamma & \epsilon \\ & \epsilon & \gamma \quad \epsilon \\ & & \bullet \\ & & \bullet \end{array}\right] \text{repeats} \left\{\begin{array}{c} B_2 \\ B_4 \\ B_6 \\ \bullet \\ \bullet \\ \bullet \end{array}\right\} = \{0\} \qquad (17)$$

where

$$\gamma = (1-2\nu)\left(1 - \frac{\sigma_o}{2\mu} - q^2\right)$$

$$\omega = 2(1-\nu)\left[1 - \frac{(1-2\nu)}{2(1-\nu)}\frac{\sigma_o}{2\mu} - q^2\right]$$

$$\epsilon = \frac{(1-2\nu)}{4\mu}\sigma_o.$$

In the first system equation (16), coupling exists both between the even terms of the auxiliary functions, ϕ_m (z, ξ) and χ_m (z, ξ), and the auxiliary functions themselves. The only coupling that exists in the second system equation (17) is between the even terms of the auxiliary function, ψ_m (z, ξ).

The solution procedure requires the determinants of the two infinite systems of equations (16), (17) to vanish for nontrivial solutions to exist. In order to make the formulation mathematically tractable, the complete coupling of the auxiliary functions arising from the first two terms of the integral-series representations equation (14) are considered, while the effect of coupling is disregarded[3] for the remaining terms. Even if the completely coupled infinite system of equations (16), (17) were solved, this analysis would not provide the buckling stress associated with the higher-order modes due to the interpenetration of delamination surfaces and symmetry requirements. However, the exact solution would account for the effect of all of the higher-order modes on the fundamental buckling stress.

This approximation leads to the partial coupling of the two systems of homogeneous algebraic equations as shown in equations (16) and (17) with bold lines. The effect of this approximation and the coupling between the even terms of the auxiliary functions are discussed in Section VI.

The characteristic equation of the reduced system shown with bold lines in equation (16) is $[(1+\gamma)(1-\omega) - q^2]^2 - 2\epsilon^2[(1+\gamma)^2 + (1-\omega)^2 - 2q^2] + 4\epsilon^4 = 0$ and it is satisfied if

$$q_{1,2} = \mp\gamma_1 \quad ; \quad \gamma_1 = \sqrt{1 - \frac{1-2\nu}{2(1-\nu)}\frac{\sigma_o}{2\mu}(1 - 1/\sqrt{2})}$$

$$q_{3,4} = \mp\gamma_2 \quad ; \quad \gamma_2 = \sqrt{1 - \frac{\sigma_o}{2\mu}(1 - 1/\sqrt{2})}$$

$$q_{5,6} = \mp\gamma_3 \quad ; \quad \gamma_3 = \sqrt{1 - \frac{1-2\nu}{2(1-\nu)}\frac{\sigma_o}{2\mu}(1 + 1/\sqrt{2})}$$

$$q_{7,8} = \mp\gamma_4 \quad ; \quad \gamma_4 = \sqrt{1 - \frac{\sigma_o}{2\mu}(1 + 1/\sqrt{2})}.$$

The roots of the characteristic equation, $\gamma = 0$, resulting from the second reduced system shown with bold lines in equation (17) are

$$q_{9,10} = \mp\gamma_5 \quad ; \quad \gamma_5 = \sqrt{1 - \frac{\sigma_o}{2\mu}}.$$

The general solutions for ϕ_o (z, ξ), χ_o (z, ξ), ϕ_2 (z, ξ), χ_2 (z, ξ), and ψ_2 (z, ξ) then become

$$\begin{Bmatrix} \phi_o(z,\xi) \\ \chi_o(z,\xi) \\ \phi_2(z,\xi) \\ \chi_2(z,\xi) \end{Bmatrix} = X_1 \begin{Bmatrix} 1/\sqrt{2} \\ \rho(\gamma_1) \\ 1 \\ \sqrt{2}\,\rho(\gamma_1) \end{Bmatrix} e^{\gamma_1 \xi z}$$

$$+ X_2 \begin{Bmatrix} 1/\sqrt{2} \\ -\rho(\gamma_1) \\ 1 \\ -\sqrt{2}\,\rho(\gamma_1) \end{Bmatrix} e^{-\gamma_1 \xi z} + X_3 \begin{Bmatrix} 1/\sqrt{2} \\ \rho(\gamma_2) \\ 1 \\ \sqrt{2}\,\rho(\gamma_2) \end{Bmatrix} e^{\gamma_2 \xi z}$$

[3]This approximation is supported by the fact that the off-diagonal terms ϵ are of order σ_o/μ and therefore small compared to the diagonal terms γ, $1+\gamma$, $1-\omega$, which are of order one.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$+ X_4 \begin{Bmatrix} 1/\sqrt{2} \\ -\rho(\gamma_2) \\ 1 \\ -\sqrt{2}\,\rho(\gamma_2) \end{Bmatrix} e^{\gamma_2 \xi z} + X_5 \begin{Bmatrix} -1/\sqrt{2} \\ \rho(\gamma_3) \\ 1 \\ -\sqrt{2}\,\rho(\gamma_3) \end{Bmatrix} e^{\gamma_3 \xi z}$$

$$+ X_6 \begin{Bmatrix} 1/\sqrt{2} \\ -\rho(\gamma_3) \\ 1 \\ \sqrt{2}\,\rho(\gamma_3) \end{Bmatrix} e^{-\gamma_3 \xi z}$$

$$+ X_7 \begin{Bmatrix} -1/\sqrt{2} \\ \rho(\gamma_4) \\ 1 \\ -\sqrt{2}\,\rho(\gamma_4) \end{Bmatrix} e^{\gamma_4 \xi z} + X_8 \begin{Bmatrix} -1/\sqrt{2} \\ -\rho(\gamma_4) \\ 1 \\ \sqrt{2}\,\rho(\gamma_4) \end{Bmatrix} e^{-\gamma_4 \xi z} \tag{18a}$$

$$\psi_2(z, \xi) = X_9\, e^{\gamma_5 \xi z} + X_{10}\, e^{-\gamma_5 \xi z} \tag{18b}$$

where

$$\rho(q) = -\frac{[(1-\gamma)^2 - q^2 - 2\epsilon]\epsilon}{[(1+\gamma)(1-\omega) - q^2 - 2\epsilon^2]q}.$$

Application of the traction-free conditions on S_1 and the shear-free conditions in the midplane of the layer permits the expression of eight of the coefficients in terms of X_4 and X_8. The remaining unknowns X_4 and X_8 are determined by imposing the mixed boundary conditions in the plane of the delamination resulting in the following two simultaneous pairs of coupled dual integral equations:

$$\int_0^\infty [k_1 S(\xi) + k_2 D(\xi)]\xi J_o(r\xi)d\xi = \int_0^\infty \left[h\left(\xi; \frac{\sigma_o}{\mu}, t\right) S(\xi) + g\left(\xi; \frac{\sigma_o}{\mu}, t\right) D(\xi) \right] \xi J_o(r\xi)d\xi$$

$$r \epsilon [0, 1) \tag{19a}$$

$$\int_0^\infty [\kappa_2 S(\xi) + \kappa_1 D(\xi)]\xi J_2(r\xi)d\xi = \int_0^\infty \left[g\left(\xi; \frac{\sigma_o}{\mu}, t\right) S(\xi) + h\left(\xi; \frac{\sigma_o}{\mu}, t\right) D(\xi) \right] \xi J_2(r\xi)d\xi$$

$$r \epsilon [0, 1) \tag{19b}$$

$$\int_0^\infty D(\xi) J_o(r\xi)d\xi = 0 \qquad r \epsilon [1, \infty) \tag{19c}$$

$$\int_0^\infty S(\xi) J_2(r\xi)d\xi = 0 \qquad r \epsilon [1, \infty) \tag{19d}$$

where $S(\xi)$ and $D(\xi)$ are unknown functions. The relationship between $S(\xi)$, $D(\xi)$ and X_4, X_8, as well as expressions for $h(\xi; \frac{\sigma_o}{\mu}, t)$, $g(\xi; \frac{\sigma_o}{\mu}, t)$ and the constants κ_1, κ_2 are presented in the Appendix.

These coupled dual integral equations (19) are converted to a pair of coupled homogeneous Fredholm integral equations of the second kind by employing the following integral representations for $S(\xi)$ and $D(\xi)$ (Westmann, 1965).

$$S(\xi) = \xi^{1/2} \int_0^1 \alpha^{1/2} \Phi_2(\alpha) J_{5/2}(\xi\alpha)d\alpha \tag{20a}$$

$$D(\xi) = \xi^{1/2} \int_0^1 \alpha^{1/2} \Phi_1(\alpha) J_{1/2}(\xi\alpha)d\alpha - \frac{\kappa_1}{\kappa_2} \xi^{1/2} \int_0^1 \alpha^{1/2} \Phi_2(\alpha) J_{5/2}(\xi\alpha)d\alpha \tag{20b}$$

in which $\Phi_1(\alpha)$, $\Phi_2(\alpha)$ are unknown functions taken to be continuous in the interval [0, 1] and are required to satisfy the conditions

$$\lim_{\alpha \to 0+} [\alpha^{1/2} \Phi_1(\alpha)] = 0, \quad \lim_{\alpha \to 0+} [\alpha^{1/2} \Phi_2(\alpha)] = 0.$$

Using the Weber-Schaftheitlin integral (Watson, 1962) it may be shown that the representations for $S(\xi)$ and $D(\xi)$ automatically satisfy equations (19*c*) and (19*d*).

Modifying $D(\xi)$ by integrating the first integral in equation (20*b*) by parts and substituting the resulting expression along with equation (20*a*) into (19*a*) lead to the following Abel integral equation

$$\kappa_2 \int_0^r \frac{d}{d\alpha} [\Phi_1(\alpha)] \frac{d\alpha}{\sqrt{r^2 - \alpha^2}} = \sqrt{\frac{\pi}{2}} \int_0^\infty [h(\xi \frac{\sigma_o}{\mu}, t) S(\xi) + g(\xi; \frac{\sigma_o}{\mu} t) D(\xi)] \xi\, J_o(r\xi)d\xi. \tag{21}$$

Again, modifying both equations (20*a*), (20*b*), by integrating by parts and substituting the resulting expressions into equation (19*b*) yield the second Abel integral equation

$$\int_0^r \left\{ \frac{(\kappa_2^2 - \kappa_1^2)}{\kappa_2} \frac{d}{d\alpha} [\alpha^2 \Phi_1(\alpha)] - \kappa_1 \alpha^3 \frac{d}{d\alpha} [\alpha^{-1} \Phi_1(\alpha)] \right\} \frac{d\alpha}{\sqrt{r^2 - \alpha^2}}$$

$$= \sqrt{\frac{\pi}{2}}\, r^2 \int_0^\infty \left[g\left(\xi; \frac{\sigma_o}{\mu}, t\right) S(\xi) + h\left(\xi; \frac{\sigma_o}{\mu}, t\right) D(\xi) \right] \xi\, J_2(r\xi)d\xi. \tag{22}$$

Applying Abel's solution method to equations (21) and (22), and substituting the expressions for $S(\xi)$, $D(\xi)$, and equations (20), into the right-hand side of equations (21), (22) result in the final integral equations for $\Phi_1(\alpha)$, $\Phi_2(\alpha)$;

$$\begin{bmatrix} k_2 & 0 \\ -k_1 & (k_2^2 - k_1^2)/k_2 \end{bmatrix} \begin{Bmatrix} \Phi_1(\alpha) \\ \Phi_2(\alpha) \end{Bmatrix} = \lambda^* \int_0^1 \begin{bmatrix} k_{11}\left(\alpha, \eta; \frac{\sigma_o}{\mu}, t\right) & k_{12}\left(\alpha, \eta; \frac{\sigma_o}{\mu}, t\right) \\ k_{21}\left(\alpha, \eta; \frac{\sigma_o}{\mu}, t\right) & k_{22}\left(\alpha, \eta; \frac{\sigma_o}{\mu}, t\right) \end{bmatrix} \begin{Bmatrix} \Phi_1(\eta) \\ \Phi_2(\eta) \end{Bmatrix} d\eta$$

$$(\lambda^* = 1). \tag{23}$$

The kernels of the coupled system of homogeneous Fredholm integral equations of the second kind equation (23) are given in terms of the following integrals

$$K_{11}(\alpha, \eta) = \sqrt{\alpha\eta} \int_0^\infty g\left(\xi; \frac{\sigma_o}{\mu}, t\right) \xi J_{1/2}(\alpha\xi) J_{1/2}(\eta\xi)d\xi$$

$$K_{12}(\alpha, \eta) = \sqrt{\alpha\eta} \int_0^\infty \left[h\left(\xi; \frac{\sigma_o}{\mu}, t\right) - \frac{\kappa_1}{\kappa_2} g\left(\xi; \frac{\sigma_o}{\mu}, t\right) \right] \xi\, J_{1/2}(\alpha\xi) J_{5/2}(\eta\xi)d\xi \tag{24}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

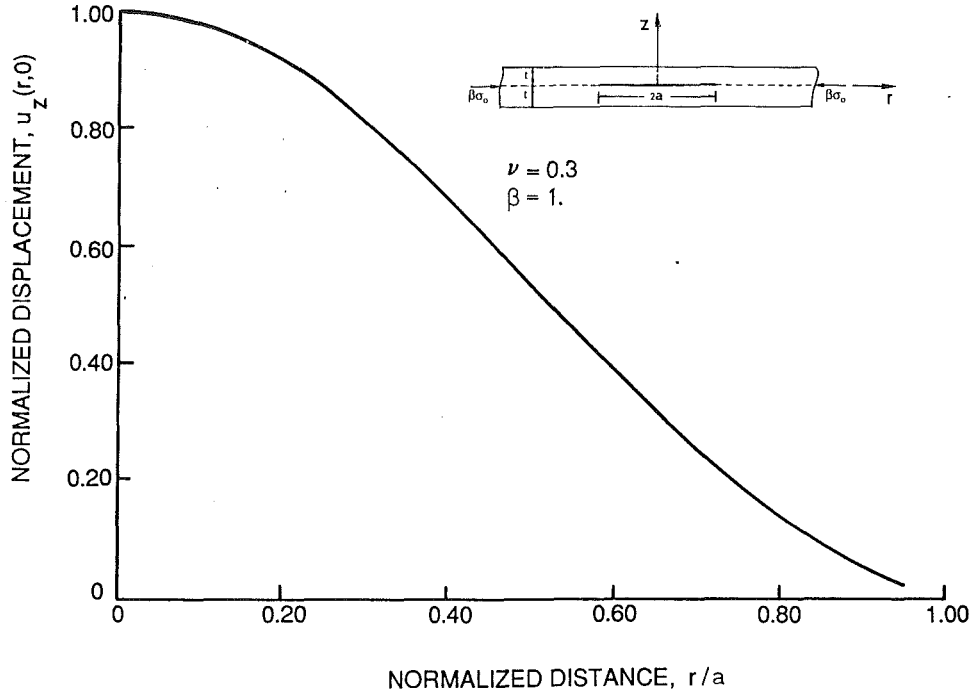


Fig. 3 Buckled delamination shape—axisymmetric load

$$K_{21}(\alpha,\eta)=\sqrt{\alpha\eta}\int_0^\infty h\left(\xi;\frac{\sigma_o}{\mu},t\right)\xi J_{5/2}(\alpha\xi)\times J_{1/2}(\eta\xi)d\xi-3k_1\eta H(\alpha-\eta)/\alpha^2$$

$$K_{22}(\alpha,\eta)=\sqrt{\alpha\eta}\int_0^\infty\left[g\left(\xi;\frac{\sigma_o}{\mu},t\right)-\frac{\kappa_1}{\kappa_2}h\left(\xi;\frac{\sigma_o}{\mu},t\right)\xi\, J_{5/2}(\alpha\xi)J_{5/2}(\eta\xi)d\xi\right.$$

where $H(x)$ is the the Heaviside step function. The solution is now completed by finding those values of σ_o so that $\lambda^* = 1$ is an eigenvalue and nontrivial solutions to equations (23) exist.

V Numerical Results

In the problems under consideration, the kernels of the homogeneous integral equations (10) and (23) contain the unknown applied compressive stress, σ_o, that causes buckling. Therefore, the solution procedure involves searching for σ_o/μ such that the $\lambda^* = 1$ is an eigenvalue in these equations.

The complexity of the kernels in equations (10), (23) requires that the integral equations be solved numerically. The well-known numerical solution procedure involves the reduction of these integral equations to a system of homogeneous algebraic equations. A nontrivial solution of the integral equation is found by searching for the values of σ_o/μ which cause the determinant of this system of algebraic equations to vanish. This scheme was used successfully by Stahl and Keer (1972) to solve the same types of integral equations.

The numerical treatment of the integral equation (10) provides the variation of the dimensionless buckling stress, σ_o/μ, as a function of the dimensionless thickness, t/a (Fig. 2).

Once the buckling stress is computed, then the buckling shape is obtained from the following expression

$$u_z(r,0)=\sqrt{2/\pi}\int_r^1\frac{\Psi(\alpha)}{\sqrt{\alpha^2-r^2}}d\alpha\, H(1-r). \tag{25}$$

It is important to note that, the asymptotic form of equation (25) is as follows:

$$u_z(r,0)=\sqrt{2/\pi}\Psi(1)\sqrt{1-r^2}+0[(1-r)] \text{ as } r\to 1-.$$

A typical normalized buckling shape is presented in Fig. 3 for dimensionless thickness $t/a = 0.1$. In all the calculations, the value of Poisson's ratio is taken to be 0.3.

In the case of uniaxial compression, the equations to be solved (equation (23)) are a pair of coupled homogeneous Fredholm integral equations of the second kind. The kernels, equation (24), of this system are also evaluated numerically.

The solution of the coupled system of integral equations is conducted by extending the method adopted for the solution of a single integral equation. The variation of the buckling stress, σ/μ_o, as a function of the dimensionless layer thickness, t/a, is shown in Fig. 4.

The buckled shape of the delamination under uniaxial loading is obtained from the following

$$u_z(r,\theta,0)=\sqrt{\frac{2}{\pi}}\left[\int_r^1\frac{\Phi_1(\alpha)d\alpha}{\sqrt{\alpha^2-r^2}}+r^2\int_r^1\frac{\alpha^{-2}\Phi_2(\alpha)d\alpha}{\sqrt{\alpha^2-r^2}}\text{Cos}(2\theta)\right]H(1-r)$$

$$-\sqrt{\frac{1}{\pi}\frac{\kappa_1}{\kappa_2}}\left[2\int_r^1\alpha^{-2}\sqrt{\alpha^2-r^2}\Phi_2(\alpha)d\alpha-r^2\int_r^1\frac{\alpha^{-2}\Phi_2(\alpha)d\alpha}{\sqrt{\alpha^2-r^2}}\right]H(1-r). \tag{26}$$

This expression for the buckled shape illustrates the effects of coupling terms arising from the uniaxial loading. The asymp-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

totic expansion of equation (26) results in the following leading term:

$$u_z(r,\theta,0)=\left[\sqrt{\frac{2}{\pi}}\{\Phi_1(1)+[\Phi_2(1)-\Phi_2'(1)]\mathrm{Cos}(2\theta)\}\right.$$
$$\left.+\sqrt{\frac{2}{\pi}\frac{\kappa_1}{\kappa_2}}\{\Phi_2(1)-3\Phi_2'(1)\}\right]$$
$$\times\sqrt{1-r^2}+0[(1-r)] \qquad \text{as } r\to 1-.$$

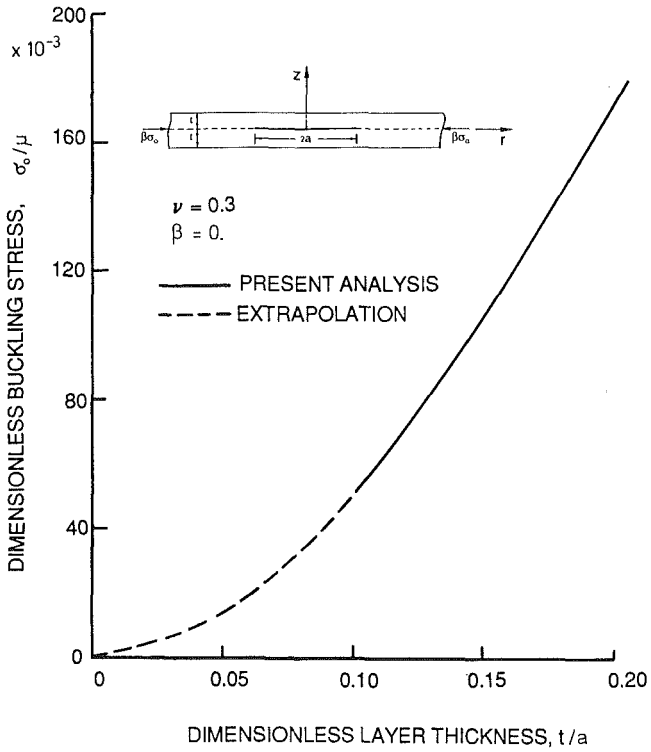


Fig. 4 Buckling stress as a function of thickness—uniaxial load

The fundamental mode shape is shown in Fig. 5 for $t/a = 0.1$ and $\theta = 0, \pi/2$.

VI Final Comments

The present analysis focuses on the determination of the buckling stress and buckled shape of the delamination under axisymmetric and uniaxial loading conditions.

Solutions of the axisymmetric loading are obtained from the treatment of a complete set of equations (1) without any approximations. For the perfect delamination and sufficiently small values of t/a in the range where the thin plate solution is valid, the buckling stress obtained from the present analysis is bounded by the results from the buckling analysis of a clamped and a simply supported, thin circular plate (Timoshenko and Gere, 1961), Fig. 2.

The buckling stress for a thin plate is proportional to the square of the plate thickness t. The dimensionless proportionality constants for the clamped and simply supported conditions are $14.68\mu/6(1-\nu)$ and $4.2\mu/6(1-\nu)$, respectively. For small values of t/a, the convergence of the numerical results becomes uneconomical because of the strong interaction between the delamination edge and bounding surfaces. However, from the converged results of the present analysis, the proportionality constant for the buckling stress for $t < 0.1$ can be approximated as $11.9/6\mu(1-\nu)$ which is shown in Fig. 2 with a dotted line.

As mentioned previously, in order to make the formulation mathematically tractable, an approximation is made leading to the partial coupling of the two systems of homogeneous algebraic equations as shown in equations (16) and (17) with bold lines. In this case, the effect of only the second mode on the fundamental buckling stress is captured.

Since the formulation deals only with the partially coupled system of equations, the significance of the coupling effect is studied by solving the following system of completely decoupled equations.

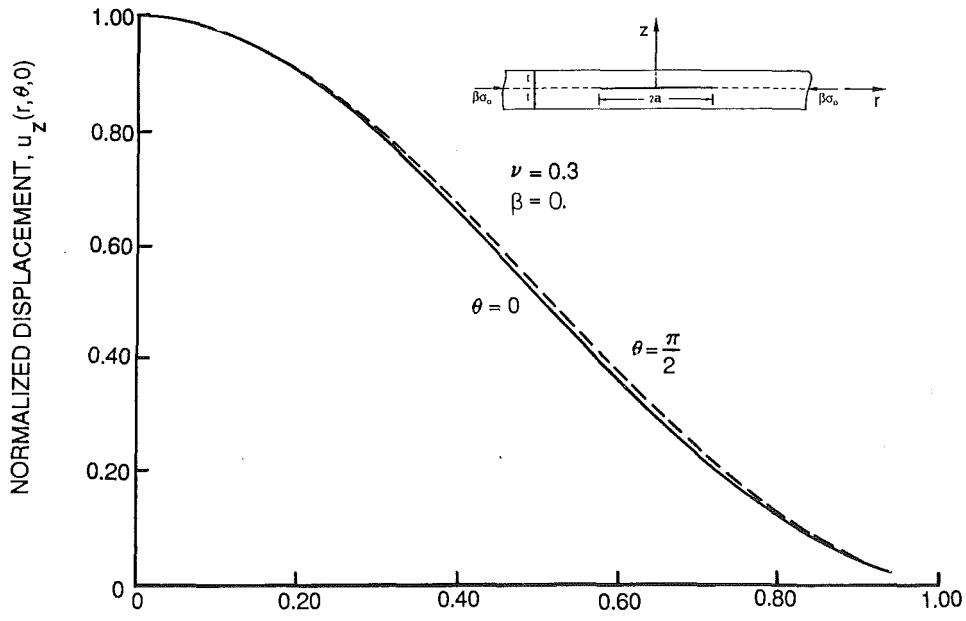


Fig. 5 Buckled delamination shape—uniaxial load

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\left[\frac{2(1-\nu)}{1-2\nu}-\frac{\sigma_o}{2\mu}\right]\phi_m\xi^2-\phi_m''+\frac{1}{1-2\nu}\chi_m'\xi=0$$

$$\left(1-\frac{\sigma_o}{2\mu}\right)\chi_m\xi^2-\frac{2(1-\nu)}{1-2\nu}\chi_m''-\frac{1}{1-2\nu}\phi_m'\xi=0 \qquad (27)$$

$$\left(1-\frac{\sigma_o}{2\mu}\right)\psi_m\xi^2-\psi_m''=0.$$

The solution of equations (27) is not acceptable for the higher modes ($m \geq 1$) due to interpenetration of delamination surfaces. This crossover is not present for the fundamental mode ($m = 0$); therefore, the solution is valid in such cases. Since equations (27) are decoupled, the solution for the fundamental buckling stress is calculated without accounting for the effect of higher-order modes.

The solution procedure for equations (27), when $m = 0$, involves operations similar to those used in the axisymmetric loading case. The form of the resulting dual integral equations is the same as in equation (7) with the exception that the definitions of τ_1 and τ_2 become

$$\tau_1=\sqrt{1-\frac{(1-2\nu)}{2(1-\nu)}\frac{\sigma_o}{2\mu}}$$

$$\tau_2=\sqrt{1-\frac{\sigma_o}{2\mu}}.$$

The fundamental buckling stress obtained from this decoupled analysis is approximately six percent higher than that from the partially coupled analysis. The six percent reduction of the fundamental buckling stress is the effect of the second mode captured in the partially decoupled analysis.

References

Bolotin, V. V., 1963, *Nonconservative Problems of the Theory of Elastic Stability*, MacMillan Publishing Co., New York.

Bottega, W. J., and Maewal, A., 1983, "Delamination Buckling and Growth in Laminates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 184–189.

Chai, H., and Babcock, C. D., 1985, "Two-Dimensional Modelling of Compressive Failure in Delaminated Laminates," *J. Comp. Materials*, Vol. 19, pp. 67–98.

Chai, H., Babcock, C. D., and Knauss, W. G., 1981, "One Dimensional Modelling of Failure in Laminated Plates by Delamination Buckling," *Int. J. Solids Structures*, Vol. 17, pp. 1069–1083.

Chai, H., Knauss, W. G., and Babcock, C. D., 1983, "Observation of Damage Growth in Compressively Loaded Laminates," *Experimental Mechanics*, Vol. 23, pp. 329–337.

Copson, E. T., 1961, "On Certain Dual Integral Equations," *Proc. Glasgow Math. Assoc.*, Vol. 5, pp. 21–24.

Flugge, W., 1972, *Tensor Analysis and Continuum Mechanics*, Springer-Verlag, New York.

Harding, J. W., and Sneddon, I. N., 1945, "The Elastic Stresses Produced by the Identation of the Plane Surface of a Semi-Infinite Elastic Body by a Rigid Punch," *Proc. Cambridge Phil. Soc.*, Vol. 41, pp. 16–26.

Kachanov, L. M., 1976, "Separation of Composite Materials," *Polymer Mechanics*, Vol. 12, pp. 812–815.

Keer, L. M., Nemat-Nasser, S., and Oranratnachai, A., 1982, "Surface Instability and Splitting in Compressed Brittle Elastic Solids Containing Crack Arrays," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 761–767.

McCarty, J. E., and Ratwani, M. M., 1983, "Damage Tolerance of Composites," F33615-82-C3212, Interim Report, No. 2, Boeing Military Airplane Company.

Muki, R., 1960, "Asymmetric Problems of the Theory of Elasticity for a Semi-Infinite Solid and a Thick Plate," *Progress in Solid Mechanics*, Vol. 1, I. N. Sneddon and R. Hill, eds., North Holland Publishing Company, Amsterdam.

Simitses, G. J., Sallam, S., and Yin, W. L., 1984, "Effect of Delamination on Axially-Loaded Laminated Plates", Proc. AIAA/ASME/ASCE/AHS 25th SDM Conference, Paper No. 84-09-64, Palm Springs, Calif.

Sneddon, I. N., and Lowengrub, M., 1969, *Crack Problems in the Classical Theory of Elasticity*, John Wiley and Sons, New York.

Stahl, B., and Keer, L. M., 1972, "Vibration and Stability of Cracked Rectangular Plates," *Int. J. Solids Structures*, Vol. 8, pp. 69–91.

Starnes, J. H., and Williams, J. G., 1984, "Failure Characteristics of Graphite/Epoxy Structural Components Loaded in Compression," NASA TM 84552.

Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.

Washizu, K., 1975, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, New York.

Watson, G. N., 1962, *Theory of Bessel Functions*, University Press, Cambridge, U.K.

Westmann, R. A., 1965, "Simultaneous Pairs of Dual Integral Equations," *Society of Industrial Applied Mathematics Review*, Vol. 7, pp. 341–348.

Whittaker, E. T., and Watson, G. N., 1920, *A Course of Modern Analysis*, University Press, Cambridge, U.K.

Yin, W. L., and Fei, Z., 1985, "Delamination Buckling and Growth in a Clamped Circular Plate,"Proc. AIAA/ASME/ASCE/AHS 26th SDM Conference, Paper No. 85-08-32, Orlando, Fla.

Yin, W. L., Sallam, S., and Simitses, G. J., 1984, "Ultimate Axial Load Capacity of a Delaminated Plate," Proc. AIAA/ASME/ASCE/AHS 25th SDM Conference, Paper No. 84-08-92, Palm Springs, CA.

Yin, W. L., and Wang, J. T. S., 1984, "The Energy Release Rate in the Growth of a One-Dimensional Delamination," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 939–941.

APPENDIX

In equation (19), the unknown functions $S(\xi)$ and $D(\xi)$ are related to the unknown coefficients X_4 and X_8 as follows:

$$\begin{aligned} S(\xi) &= [(k_{14}-k_{24})\rho(\gamma_1)+(k_{34}-1)\rho(\gamma_2)]X_4 \\ &\quad -[(k_{58}-k_{68})\rho(\gamma_3)+(k_{78}-1)\rho(\gamma_4)]X_8 \qquad (A1) \\ D(\xi) &= -[(k_{14}-k_{24})\rho(\gamma_1)+k_{34}-1)\rho(\gamma_2)]X_4 \\ &\quad -[(k_{58}-k_{68})\rho(\gamma_3)+(k_{78}-1)\rho(\gamma_4)]X_8 \end{aligned}$$

where

$$k_{14}=[(b_3-b_4)e^{-(\gamma_1+\gamma_2)\xi t}+(b_3+b_4)e^{-(\gamma_1-\gamma_2)\xi t}-2b_3]/\Delta_1(\xi)$$

$$k_{24}=[(b_3-b_4)e^{(\gamma_1+\gamma_2)\xi t}+(b_3+b_4)e^{(\gamma_1-\gamma_2)\xi t}-2b_3]/\Delta_1(\xi)$$

$$k_{34}=[(b_1-b_2)e^{-(\gamma_1+\gamma_2)\xi t}+(b_1+b_2)e^{(\gamma_1-\gamma_2)\xi t}-2b_1]/\Delta_1(\xi)$$

$$k_{58}=[(c_3-c_4)e^{-(\gamma_3+\gamma_4)\xi t}+(c_3+c_4)e^{-(\gamma_3-\gamma_4)\xi t}-2c_3]/\Delta_2(\xi)$$

$$k_{68}=[(c_3-c_4)e^{(\gamma_3+\gamma_4)\xi t}+(c_3+c_4)e^{(\gamma_3-\gamma_4)\xi t}-2c_3]/\Delta_2(\xi)$$

$$k_{78}=[(c_1-c_2)e^{-(\gamma_3+\gamma_4)\xi t}+(c_1+c_2)e^{(\gamma_3-\gamma_4)\xi t}-2c_1]/\Delta_2(\xi) \qquad (A2)$$

in which

$$\Delta_1(\xi)=(b_1-b_2)e^{(\gamma_1+\gamma_2)\xi t}+(b_1+b_2)e^{-(\gamma_1-\gamma_2)\xi t}-2b_1$$

$$\Delta_2(\xi)=(c_1-c_2)e^{(\gamma_3+\gamma_4)\xi t}+(c_1+c_2)e^{-(\gamma_3-\gamma_4)\xi t}-2c_1.$$

The constants, which appear in equation (A2), are defined as follows:

$$b_1=a_{11}\,a_{13}\,a_{31},\ b_2=a_{11}^2a_{33},\ b_3=a_{11}\,a_{13}\,a_{33},\ b_4=a_{13}^2a_{31},$$

$$c_1=a_{15}\,a_{17}\,a_{35},\ c_2=a_{15}^2a_{37},\ c_3=a_{15}\,a_{17}\,a_{37},\ c_4=a_{17}^2\,a_{35}$$

where

$$a_{11}=\gamma_1/\sqrt{2}-\rho(\gamma_1) \qquad a_{31}=\frac{2}{(1-2\nu)}[(1-\nu)\gamma_1\rho(\gamma_1)+\nu/\sqrt{2}]$$

$$a_{13}=\gamma_2/\sqrt{2}-\rho(\gamma_2) \qquad a_{33}=\frac{2}{(1-2\nu)}[(1-\nu)\gamma_2\rho(\gamma_2)+\nu/\sqrt{2}]$$

$$a_{15}=-\gamma_3/\sqrt{2}-\rho(\gamma_3) \qquad a_{35}=\frac{2}{(1-2\nu)}[(1-\nu)\gamma_3(\gamma_3)-\nu/\sqrt{2}]$$

$$a_{17}=-\gamma_4/\sqrt{2}-\rho(\gamma_4) \qquad a_{37}=\frac{2}{(1-2\nu)}[(1-\nu)\gamma_4(\gamma_4)-\nu/\sqrt{2}].$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The function $\rho(q)$ and the roots $\gamma_1 - \gamma_4$ are given in Section IV.

The known functions $h\left(\xi; \frac{\sigma_o}{\mu}, t\right)$ and $g\left(\xi; \frac{\sigma_o}{\mu}, t\right)$ are defined as

$$h\left(\xi; \frac{\sigma_o}{\mu}, t\right) = \kappa_1 - \left[H_2\left(\xi; \frac{\sigma_o}{\mu}, t\right) - H_1\left(\xi; \frac{\sigma_o}{\mu}, t\right)\right]$$

$$g\left(\xi; \frac{\sigma_o}{\mu}, t\right) = \kappa_2 - \left[H_2\left(\xi; \frac{\sigma_o}{\mu}, t\right) + H_1\left(\xi; \frac{\sigma_o}{\mu}, t\right)\right]$$

where

$$H_1(\xi) = \frac{(k_{14} + k_{24})a_{31} + (k_{34} + 1)a_{33}}{(k_{14} - k_{24})\rho(\gamma_1) + (k_{34} - 1)\rho(\gamma_2)}$$

$$H_2(\xi) = \frac{(k_{58} + k_{68})a_{35} + (k_{78} + 1)a_{37}}{(k_{58} - k_{68})\rho(\gamma_3) + (k_{78} - 1)\rho(\gamma_4)}.$$

The constants k_1 and k_2 are defined as

$$\frac{1}{2}(k_2 + k_1) = \frac{a_{11}\, a_{33} - a_{13}\, a_{31}}{a_{13}\rho(\gamma_1) - a_{11}\rho(\gamma_2)}$$

$$\frac{1}{2}(k_2 - k_1) = \frac{a_{15}\, a_{37} - a_{17}\, a_{35}}{a_{17}\rho(\gamma_3) - a_{15}\rho(\gamma_4)}.$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Buckling Calculations and Measurements on a Technologically Relevant Toroidal Magnet System

W. Geiger
Institut für Datenverarbeitung
in der Technik

K. P. Jungst
Institut für Technische Physik

Kernforschungszentrum Karlsruhe,
D-7500 Karlsruhe,
Federal Republic of Germany

The coils in a toroidal magnet system may deform laterally and become mechanically unstable when a critical value of the current is exceeded as a result of magnetoelastic coupling. This instability is called magnetoelastic buckling. It was studied in the TESPE arrangement, a toroidal magnet system the design and size of which meet technical demands. A generalized theoretical model was developed for the calculation of the buckling behavior of toroidal magnet systems. It was shown by the measurements that in a torus of complex design, the lateral elastic stiffnesses of the coils may vary considerably with the current. Monitoring of the natural frequencies of the vibration modes proved to be a suitable method for the supervision of the mechanical stability during operation.

Introduction

In a toroidal coil arrangement as shown in Fig. 1, the azimuthal components of the magnetic forces between the coils cancel each other in case of exact symmetry. In case of a small lateral deflection of a coil out of the symmetrical position, however, this coil experiences a resultant azimuthal force which tends to increase the deflection. This force is proportional both to the displacement itself and to the square of the coil current. The effect of this force can be described by a negative magnetic stiffness proportional to the square of the current. The negative magnetic stiffness is superimposed on the positive elastic stiffness of the mechanical structure which tries to keep the coil at its original position. With an increasing coil current, total stiffness of the lateral deflection is reduced. In case a critical current value is exceeded, total stiffness turns negative, i.e., the magnet arrangement becomes mechanically unstable. This effect is called magnetoelastic buckling. It is of special importance to the mechanical stability of the toroidal field magnet system of large tokamak fusion assemblies as well as of a new generation of particle detectors, namely, the large acceptance spectrometers with a toroidal field as proposed for the CEBAF accelerator. Detectable effects of buckling can be noticed already at currents below the critical value:

(1) The frequencies of the lateral vibration modes of the coils decrease with increasing current.

(2) In case of a misalignment at zero current, affected coils experience an increasing lateral deflection.

Up to now, the buckling effect in toroidal magnet systems has been experimentally investigated only by F. C. Moon and co-workers at Cornell University (Moon and Swanson, 1977; Moon, 1979, 1984) and by K. Miya and co-workers at the University of Tokyo (Miya et al., 1982; Miya and Uesaka, 1982). In these experiments, however, the model coils used were small and flexible and had no casing. They had a diameter of 0.12–0.2 m and maximum current amounted to about 10^4 ampere-turns. With these experiments, the effect could be basically demonstrated and the corresponding theory was developed. But it was necessary to fill in the large gap between those small laboratory coils and the coils of a reactor-relevant plasma experiment like NET, as there are differences of two orders of magnitude in size, three orders of magnitude in ampere-turns, and several orders of magnitude in lateral

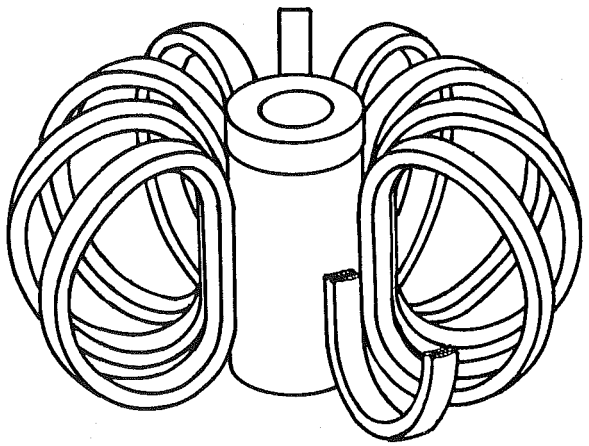


Fig. 1 Schematic drawing of the toroidal magnet system of a tokamak

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 10, 1989; final revision, Oct. 27, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Main parameters of TESPE

Number of *D*-shaped coils	6
Number of windings per coil	240
Operating current	7000 A
Maximum magnetic field at 7 kA	7 T
Stored energy	8.3 MJ
Mass of one coil	700 kg
Total mass of intercoil structure	2900 kg
Outer coil diameter (vertical)	1.05 m
Outer torus diameter	2.55 m

elastic stiffness. The TESPE toroidal magnet system seemed to be very suitable for this purpose. TESPE is a torus with 6 D-shaped superconducting coils of technologically relevant size and construction. The outer coil diameter is about 1 m. Under operating conditions, the current in each coil amounts to 1.7×10^6 ampere-turns. The windings are embedded into a massive steel casing.

For a toroidal magnet system the relevance of buckling depends on the amount of structural material which can be accepted. In case of a rigid azimuthal intercoil structure, e.g., designed to withstand the high asymmetric forces when zero current occurs in one coil and maximum current in the others, buckling effects certainly are of no concern. However, in large future reactor systems the aforementioned requirement may come into conflict with the needed "transparent" structure giving way, e.g., for plasma heating or remote handling systems. The attempt to reduce the structure to meet the latter requirement may lead to a situation where the buckling effect becomes the limiting factor.

The first step of the experimental program consisted of buckling investigations using a special test coil designed for low buckling current. It was the objective of this part of the program to confirm experimentally the calculation methods applied and to check the measurement and evaluation procedures. The second step included the calculation of the buckling behavior of the TESPE torus on the basis of the relevant deformation modes and the related elastic and magnetic stiffnesses of the magnet assembly. The deformation modes and elastic stiffnesses were determined by an experimental modal analysis at zero current and the magnetic stiffnesses were obtained by numerical calculation. The TESPE investigations showed that the TESPE assembly is described in a somewhat inaccurate manner by the model used by Moon for the calculation of the buckling behavior of toroidal magnet systems. Therefore, a more general model was developed and applied. The third step finally consisted of the measurement of the buckling behavior of the TESPE coils up to operating current.

The TESPE Superconducting Toroidal Magnet System

The TESPE toroidal magnet system was developed to obtain experience in design and construction as well as in torus operation under normal and under off-normal conditions (Jüngst, 1984; Jüngst et al., 1985). Here, there was the opportunity to extend the experimental investigations of magnetoelastic buckling with small coils mentioned previously by use of a technology relevant magnet system. In Table 1 some TESPE parameters are listed. The coil casings are made of austenitic steel. Their massiveness allows small coil deformations only under the occurring high magnetic forces. (At operating current, each coil experiences a magnetic force of 1.3×10^6 N which is directed towards the center of torus.) Moreover, the casing has to withstand the pressure rise resulting from the rapid evaporation of the liquid helium in case of loss of vacuum.

The massive intercoil structure (cf., Fig. 2) was designed to maintain a good mechanical stability in case of zero current in one coil and full current in the others. However, small gaps were left between the segmental parts of the intercoil structure to avoid high thermal stresses resulting from differences in the contraction properties of the torus elements during cooldown. Within the limits of these gaps, the coils may deflect and vibrate laterally.

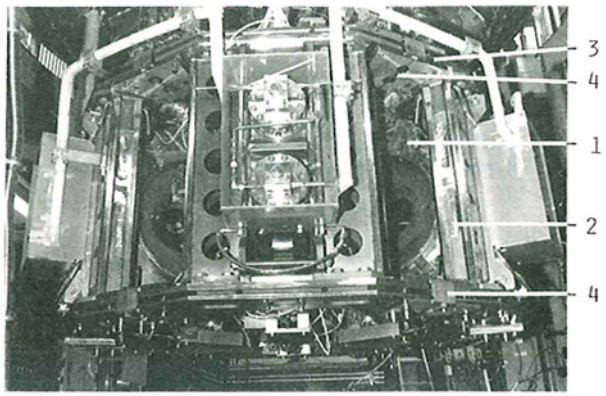


Fig. 2 TESPE torus: 1—coil, 2 and 3—vertical and segmental part of the azimuthal intercoil structure, 4—gaps in the intercoil structure

Precursor Experiments

As a first step, buckling investigations were carried out using a test coil of simple structure and low buckling current especially designed for this purpose. The background magnetic field was supplied by the TESPE torus. It was the objective of these studies to verify experimentally the numerical calculation methods and to further develop the methods of measuring and evaluating buckling effects.

Experimental. The buckling-special test coil (BS coil) was mounted on a free coil position in the symmetric three-coil arrangement of TESPE. It was installed with a deliberately chosen small angular misalignment. The BS coil was designed to enable both the exact calculation and the exact measurement of the buckling behavior—here, the decrease in lateral vibration frequency and the increase in lateral deflection with increasing current. The BS coil has 16 circular superconducting windings of 600 mm in diameter which are mounted on a rigid copper ring. The coil is mounted to the torus by means of a plate-spring which allows the coil to rotationally deflect around the plate-spring.

Theory and Numerical Calculations for the BS Coil. Let K^e be the elastic stiffness, L_2 the differential inductance, I^{BS} the current in the BS coil and I^{TS} the current in the TESPE coils. Then, the total stiffness of the coil's lateral deflection becomes zero when the product of currents I^{BS} and I^{TS} reaches K^e/L_2 (Geiger and Jüngst, 1987):

$$(I^{BS} I^{TS})_B = \frac{K^e}{L_2} . \tag{1}$$

This product is called the "buckling current product." For each given TESPE current, I^{TS}, the buckling current of the BS coil, I_B^{BS}, is obtained using equation (1):

$$I_B^{BS} = \frac{K^e}{L_2 I^{TS}} . \tag{2}$$

If ϕ_0 is the angle of misalignment at zero current, the static deflection of the coil is given by

$$\phi(I^{BS}) = \frac{\phi_0}{\frac{I_B^{BS}}{I^{BS}} - 1} . \tag{3}$$

The natural frequency f_0 (I^{BS}) of the coil's lateral vibration is obtained as

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

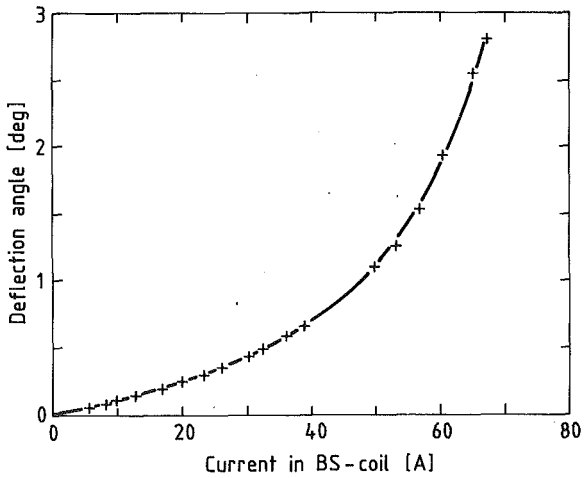


Fig. 3 Deflection ϕ of the BS coil versus current I^{BS}: measured values (+) and theoretical function equation (3) with fitted parameters (——)

$$f_0(I^{BS}) = f_0(0)\sqrt{1 - \frac{I^{BS}}{I_B^{BS}}} \tag{4}$$

with the value of the natural frequency at zero current being

$$f_0(0) = \sqrt{K^e/\Theta}/2\pi. \tag{5}$$

Θ denotes the moment of inertia with respect to the axis of rotation.

The buckling current product is well suited for an experimental verification of the theory. The "theoretical" buckling current product was determined by the calculation of L_2 using a magnetic field computer program and by a separate measurement of the elastic stiffness K^e (by measuring the deflection caused by an impressed mechanical torque). Corresponding values of $L_2 = 6.57 \times 10^{-4}$ H/rad^2 and $K^e = 263$ Nm/rad were obtained. Using equation (1), a buckling current product of 4.00×10^5 A^2 was determined. Angular misalignment of the coil was measured geometrically and amounted to 0.83 deg.

Measurements. The static deflections of the BS coil at a TESPE current of $I^{TS} = 4520$ A are shown in Fig. 3. The measured deflections may be evaluated in form of a Southwell plot as done by Moon and Swanson (1977) and also by Geiger and Jüngst (1987). Here, another method for determination of the buckling current from static deflections is used by fitting the theoretical function (3) to the measured values. The resulting fit parameters are a buckling current of $I_B^{BS} = 86.0$ A, and an angular misalignment at zero current of $\phi_0 = 0.81$ deg. Thus, the buckling current product amounts to 3.89×10^5 A^2, which is in very good agreement with the theoretical value above. In Fig. 4 the frequency-current dispersion curve obtained in a series of vibration measurements at $I^{TS} = 3180$ A is plotted. The experimental results are in good agreement with the predicted linear decrease of the frequency squared with the current of the BS coil (equation (4) squared). A linear fit was used to obtain the buckling current of $I_B^{BS} = 134$ A for a TESPE current of $I^{TS} = 3180$ A. This results in a buckling current product of 4.25×10^5 A^2, which again is in good agreement with the theoretical value above.

Summarizing, it may be concluded that these and additional experiments carried out using the BS coil (Geiger, 1987) have both verified the qualitative predictions of the buckling theory and confirmed the numerical results with high precision.

Theory

In the previous section, a system with only one degree-of-freedom was considered. Now, the general theory for a system

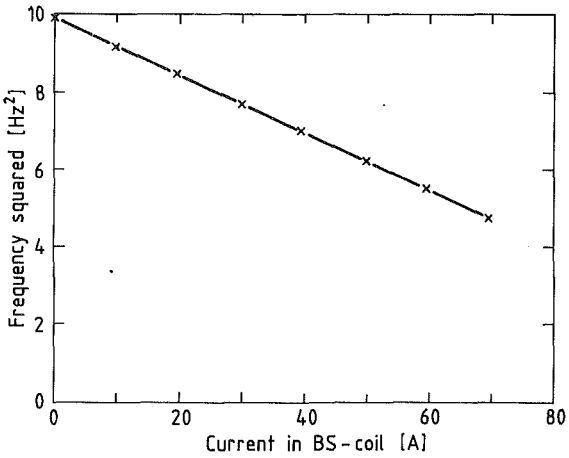


Fig. 4 Frequency-current dispersion curve, $f_0^2(I^{BS})$

of n degrees-of-freedom shall be presented, and a model for a toroidal magnet system shall be treated.

General Theory. The general theory for arbitrarily arranged coil systems (Chattopadhyay, 1979; Moon, 1980, 1984) presented here in matrix form describes a structure of N coupled elastic coils. Mechanical stability and vibrations of the coils are investigated. The system is assumed to have n mechanical degrees-of-freedom. The status of the system deformation is described uniquely by n generalized coordinates $u_i(t)$ which form the vector $\mathbf{u}(t) := (u_1(t), \ldots, u_n(t))^T$. The system is supposed to be holonomic and scleronomic and to have a stable equilibrium position of $\mathbf{u} = \mathbf{0}$, $\dot{\mathbf{u}} = \mathbf{0}$ at zero current. It is assumed that the kinetic energy is independent of the coordinates and that the current is not affected by deformations, which is true for the usual current-stabilized operation of coils. Dampings are neglected.

Within the limits of the linear theory of elasticity, the elastic potential energy and the kinetic energy of the system then can be described by the quadratic forms

$$E^{pot} = \frac{1}{2}\mathbf{u}^T\mathbf{K}^e\mathbf{u} \tag{6}$$

and

$$E^{kin} = \frac{1}{2}\dot{\mathbf{u}}^T\mathbf{M}\dot{\mathbf{u}}. \tag{7}$$

$\mathbf{K}^e$ denotes the elastic stiffness matrix, $\mathbf{M}$ the inertia matrix (mass matrix), and $\dot{\mathbf{u}}$ the generalized velocity vector of the system. The magnetic energy can be determined on the basis of the self-inductances and mutual inductances $L_{rs}(\mathbf{u})$ of the coils ($r, s = 1, 2, \ldots, N$) and the currents I_r in the coils:

$$E^{mag} = \frac{1}{2}\sum_r\sum_s L_{rs}(\mathbf{u})I_rI_s \tag{8}$$

with the double summation being taken over all coils.

The equation of motion of the system can be derived using the Lagrangian for the electromechanical system of $\mathfrak{L} = E^{kin} - E^{pot} + E^{mag}$ and Lagrange's equation of

$$\frac{d}{dt}\frac{\partial\mathfrak{L}}{\partial\dot{\mathbf{u}}} - \frac{\partial\mathfrak{L}}{\partial\mathbf{u}} = \mathbf{F}^a \tag{9}$$

where $\mathbf{F}^a$ denotes the generalized external force acting on the system. With a generalized magnetic force of

$$\mathbf{F}^m := \frac{\partial E^{mag}}{\partial\mathbf{u}} = \frac{1}{2}\sum_r\sum_s I_rI_s\frac{\partial}{\partial\mathbf{u}}L_{rs}(\mathbf{u}), \tag{10}$$

the following equation of motion is obtained:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}^e\mathbf{u} - \mathbf{F}^m = \mathbf{F}^a. \tag{11}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

By expanding the inductances $L_{rs}(\mathbf{u})$ in equation (10) in a series of $\mathbf{u}$, and stopping the expansion after the quadratic term,

$$\mathbf{F}^m = \mathbf{F}_0^m - \mathbf{K}^m \mathbf{u} \tag{12}$$

is obtained with

$$\mathbf{F}_0^m := -\frac{1}{2} \sum_r \sum_s I_r I_s \frac{\partial}{\partial \mathbf{u}} L_{rs}(\mathbf{u}) \Big|_{\mathbf{u}=\mathbf{0}} \tag{13}$$

and

$$\mathbf{K}^m := -\frac{1}{2} \sum_r \sum_s I_r I_s \frac{\partial^2}{\partial \mathbf{u}^2} L_{rs}(\mathbf{u}) \Big|_{\mathbf{u}=\mathbf{0}} . \tag{14}$$

$\mathbf{F}_0^m$ is the generalized magnetic force at $\mathbf{u}=\mathbf{0}$. $\mathbf{K}^m$ is called the magnetic stiffness matrix the coefficients of which depend on the currents. $\partial^2/\partial \mathbf{u}^2$ is the Hesse matrix.

Using equation (12), the equation of motion is obtained from equation (11) in linear approximation:

$$\mathbf{M}\ddot{\mathbf{u}} + (\mathbf{K}^e + \mathbf{K}^m)\mathbf{u} - \mathbf{F}_0^m = \mathbf{F}^a. \tag{15}$$

With increasing currents the system becomes magnetoelastically indifferent or even unstable when the generalized stiffness of one mode becomes zero or negative, respectively. Zero stiffness is reached when one natural frequency becomes zero. The natural frequencies can be derived from the equation of motion (15) using modal analysis. This leads to the general matrix eigenvalue problem

$$(-\omega_0^2 \mathbf{M} + (\mathbf{K}^e + \mathbf{K}^m))\hat{\mathbf{u}}_0 = \mathbf{0} \tag{16}$$

where the ω_{0i} denote the natural frequencies and $\hat{\mathbf{u}}_{0i}$ the eigenvectors describing the shapes of the vibration modes. Equation (16) is the frequency-current dispersion equation. At least one frequency becomes zero, if $\det(\mathbf{K}^e + \mathbf{K}^m) = 0$.

Model for a Toroidal Magnet Arrangement. Force-deflection measurements on TESPE coils have shown that the stiffnesses of the lateral deflections of the six coils vary considerably from coil to coil. Moreover, the coils are elastically coupled by the supply lines (superconducting current joints, helium lines) which again have different stiffnesses. This elastic behavior of the TESPE torus cannot be described precisely enough by the model of toroidal magnet arrangements used by Moon for the evaluation of his torus experiments because of the simplifying assumptions, such as equal elastic stiffnesses of the coils and equal or no elastic coupling between the coils. Moreover, it is shown that the assumption made by Moon (1979, 1984) that the forces between the coils only depend on the differences between the coil deflections is not true for the TESPE torus. This was found out by the calculation of the magnetic forces for the overturning mode. In case of in-phase deflections of the coils in this mode, the forces between the coils are not negligibly small and have a significant influence on the buckling current value (see below). For this reason, a more generalized model of toroidal magnet arrangements was developed for the TESPE experiments. In this model, the assumptions made with respect to elastic and magnetic stiffnesses are less simplifying.

The following assumptions are made: The toroidal magnet system consists of N coils. Each coil has one mode of lateral elastic deflection only. The mode shape is the same for all coils and is not changed by magnetic forces. The generalized coordinate u_i describes the amplitude of deflection of the coil i out of its equilibrium position at zero current. The same current flows through all coils. The coils are arranged symmetrically (except for small misalignments). The assumptions of the general theory for arbitrarily arranged coil systems presented above are fulfilled. As far as the elastic stiffnesses are concerned, no simplifying assumptions are made, e.g., the coil stiffnesses may vary from coil to coil, and the coils may be coupled by different stiffnesses.

On the basis of the general theory, the equation of motion (15) is obtained for this system. Now, the matrices of this equation are $N \times N$ matrices and the vectors have N components each.

Magnetic Stiffness Matrix. In equation (14), the magnetic stiffness matrix is determined in the equilibrium position at zero current. At zero current there are only small coil misalignments relative to the symmetrical arrangement as a result of which the magnetic stiffness matrix is changed only very little. Therefore, the assumption of a symmetrical arrangement is a good approximation for the determination of the magnetic stiffness matrix. In this arrangement, the magnetic properties of the system are supposed to be N-fold rotationally symmetric and mirror symmetric.

By expanding the inductances $L_{rs}(u_r, u_s)$ in a Taylor series in $(u_r - u_s)$ and $(u_r + u_s)$ and taking into account that $L_{rs}(u_r, u_s) = L_{rs}(-u_s, -u_r)$ as a result of the mirror symmetry,

$$L_{rs}(u_r, u_s) = L_0^{rs} + L_1^{rs}(u_r - u_s) + L_{2-}^{rs}(u_r - u_s)^2 + L_{2+}^{rs}(u_r + u_s)^2 \tag{17}$$

is obtained. Here, the terms of third and higher order are neglected. The coefficient L_{2-}^{se} with $r \neq s$ denotes the differential mutual inductance for antiphase deflection of coils r and s, i.e., when coils r and s deflect with the same amplitude in the *opposite* direction, whereas the coefficient L_{2+}^{rs} indicates the differential mutual inductance for in-phase deflections, i.e., when the coils deflect with the same amplitude in the *same* direction. (These coefficients may be calculated with equation (17) from the mutual inductances of coils r and s for various in-phase and antiphase deflections.) The term with the coefficient L_{2-}^{rs} does not exist for $r=s$. By inserting equation (17) into equation (14) and after having carried out some intermediate calculations, the elements of the magnetic stiffness matrix result in

$$K_{ik}^m = \begin{cases} 2I^2(L_{2-}^{ik} - L_{2+}^{ik}) & i \neq k \\ -2I^2\left[\sum_{\substack{r=1 \\ r\neq i}}^{N} (L_{2-}^{ir} + L_{2+}^{ir}) + 2L_{2+}^{ii}\right] & i = k \end{cases} \tag{18}$$

for the same current I in all coils. The "neighboring distance" of two coils r and s is defined as $q(r,s) :=$ minimum $(|r-s|, N - |r-s|)$, i.e., $q=1$ for two coils which are nearest neighbors, $q=2$ for coils which are next nearest, etc. The neighboring distance of a coil to itself is defined in the same way as $q(r,r) := 0$. As a result of the symmetry, all coil pairs with the same value of q have the same differential inductances. With c_{q-} being the antiphase and c_{q+} the in-phase differential inductance at the distance of q ($c_{q-} := L_{2-}^{r,r\pm q}$, $c_{q+} := L_{2+}^{r,r\pm q}$), equation (18) can be rewritten as

$$K_{ik}^m = \begin{cases} 2I^2(c_{q-} - c_{q+}) & i \neq k \\ -2I^2\left(\sum c_r + 2c_{0+}\right) & i = k \end{cases} \tag{19}$$

where $q = q(i,k)$ and Σc_r is defined as

$$\sum c_r := \begin{cases} 2\sum_{r=1}^{N/2-1} (c_{r-} + c_{r+}) + c_{N/2-} + c_{N/2+} & N \text{ even} \\ 2\sum_{r=1}^{(N-1)/2} (c_{r-} + c_{r+}) & N \text{ odd.} \end{cases} \tag{20}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

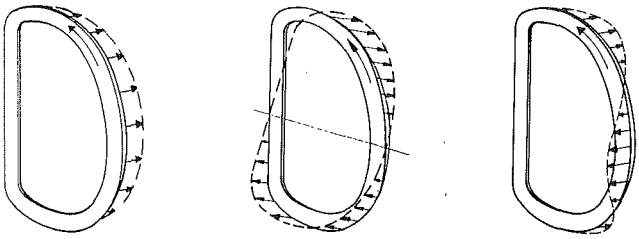


Fig. 5 Sketch of the basic mode (*a*), overturning mode (*b*), and bending mode (*c*)

Calculation of the TESPE Buckling Current

Survey of the Modes. For the calculation of the buckling current, the mode shapes and the elastic stiffnesses of the coils are required. The following results are obtained by an experimental analysis of the relevant vibration modes of the TESPE magnet system at zero current:

(1) Similar to the BS coil described previously, the TESPE coils may rotate as rigid bodies around the axis of their fixture to the central support (bucking post) as indicated in Fig. 5(*a*). Thereby, mainly the bucking post part is deformed to which the coil is attached ("front part"). In the following sections, this lateral deflection mode of the coils shall be referred to as the basic mode. The coils are coupled elastically by the supply lines and the central support structure. Thus, coupled vibrations of the coils arise within the torus compound (torus modes). The natural frequencies of the torus modes of the basic vibration vary between 13 and 18 Hz.

(2) The coils may rotate as rigid bodies around the horizontal center line (axis of symmetry in radial direction) of the coils as shown in Fig. 5(*b*). Here, deformation mainly occurs in the fixture of the coil onto the front part and in the front part itself. This deflection mode is called overturning mode. The torus modes of this coil vibration have frequencies ranging between 67 and 86 Hz.

(3) The coils may bend as indicated in Fig. 5(*c*). This deformation mode is termed bending mode. The frequencies of the corresponding torus vibration modes vary between 480 and 650 Hz.

(4) Bending modes of higher order may occur. The corresponding frequencies have values higher than 1.5 kHz.

Out of these lateral deflections, the basic mode is that deflection with the lowest buckling current. It only occurs when the radially outer portion of the coil is not fixed by the azimuthal intercoil structure of the torus. If such a fixture exists in the middle of the coil, the overturning mode is the deflection with the lowest buckling current. In case the coil is not fixed in the middle but at the two outer node lines of the bending mode, the bending mode is the deflection relevant for buckling. By investigating the different modes, the buckling currents for different intercoil structures can be estimated. The basic and the overturning modes will be analyzed; for bending mode see Geiger (1987).

Basic Mode. The TESPE investigations concentrated on the basic mode of the coils, as in this mode the calculated magnetic stiffness at design current almost equals the elastic stiffness measured under zero current conditions and detectable buckling effects could have been expected.

The calculations of the buckling behavior in the basic mode were based on the model for toroidal magnet arrangements described above. The basic deflection of the coil i was modelled as a rotation of the coil i around the axis of its fixture to the central support ($i=1, 2, \ldots, 6$). The angle ϕ_i which represents angular deflection relative to the position at zero current is chosen as the coordinate of deflection. With the vector of the coil deflections being $\boldsymbol{\phi} := (\phi_1, \ldots, \phi_6)$, the vector of the torques acting on the coils at $\boldsymbol{\phi} = \mathbf{0}$ as a result of the magnetic forces being $\boldsymbol{\tau}_0^m$ and the vector of the external torques acting on the coils being $\boldsymbol{\tau}^a$, the equation of motion (15) for this model has the form

$$\mathbf{M}\ddot{\boldsymbol{\phi}} + (\mathbf{K}^e + \mathbf{K}^m)\boldsymbol{\phi} - \boldsymbol{\tau}_0^m = \boldsymbol{\tau}^a. \tag{21}$$

Assuming that the elastic stiffnesses are independent of the current, the calculation of the buckling behavior of the basic mode at low temperatures was carried out in the following steps:

(1) The inertia matrix was calculated on the basis of geometry and mass data. The elastic stiffness matrix was experimentally determined and verified at room temperature. According to the change of Young's modulus, it was then converted to operating temperature (4 K).

(2) The magnetic stiffness matrix was determined numerically.

(3) These system matrices were used to calculate the current dependence of the natural frequencies and the buckling current.

In the following sections, these steps are described in detail:

(*1*) *Inertia Matrix and Elastic Stiffness Matrix.* In good approximation, the inertia matrix $\mathbf{M}$ is diagonal, as the mass of the supply lines is negligibly small compared with the mass of the coils. The values of the diagonal elements equal the moments of inertia of the coils around their axis of fixture. The elastic stiffness matrix at room temperature was determined experimentally by force-deflection measurements at zero current. In this case one obtains from equation (21) the equation $\mathbf{K}^e\boldsymbol{\phi} = \boldsymbol{\tau}^a$. For determination of the elastic stiffness matrix, minimum N linearly independent separate measurements are required (N=number of coils). On the TESPE torus six measurements were performed by pressing successively every two adjacent coils apart and measuring each time the force (torque) and the deflections of the coils. As these six measurements are linearly dependent, one measurement was added with an external force acting on one coil. The stiffness matrix for the basic mode was then calculated by a least square fit to the measured values and was made symmetrical by taking the average of the corresponding off-diagonal elements.

To verify these system matrices, they were applied for the calculation of the torus modes of basic vibration and the results were compared with the corresponding measured values. The torus modes at zero current were determined by solving the matrix eigenvalue problem as given by equation (16) with $\mathbf{K}^m = \mathbf{0}$ (and $\hat{\mathbf{u}}_0 \equiv \hat{\boldsymbol{\phi}}_0$). The resulting values of ω_{0i} and $\hat{\boldsymbol{\phi}}_{0i}$ represent the angular frequencies and mode shapes of the torus modes, respectively. For experimental modal analysis the vibrations were excited by a mechanical impact and the vibration signals were analyzed using a Fast Fourier Transform algorithm. Four of the six modes could be clearly identified in the vibration spectra. The measured natural frequencies agreed with the calculated values within a range of ± 3 percent. Two vibration modes could not be identified with sufficient precision and confidence from the measurements, as their frequencies were too close to another natural frequency.

For the calculation of the buckling current and the frequency-current dispersion curves, the inertia matrix and the elastic stiffness matrix at low temperatures are required. The change of the moment of inertia of the coils as a result of the cooling process is negligibly small. Thus, the room temperature inertia matrix is applicable without having to be modified. The measurement of the elastic stiffness matrix of the system at low temperatures is very expensive. For this reason, the elastic stiffness matrix at room temperature was converted to low temperatures according to the change of Young's modulus.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

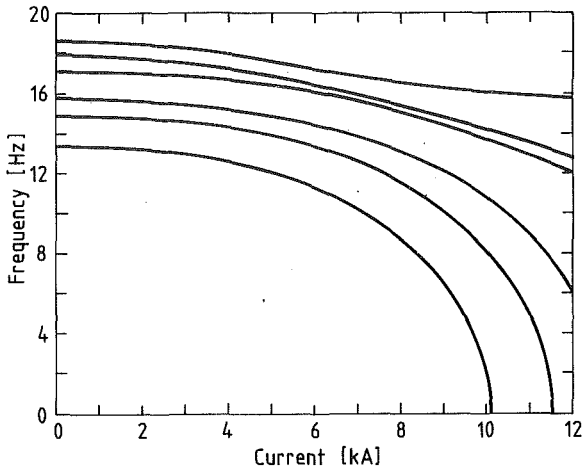


Fig. 6 Calculated frequency-current dispersion curves $f_{0i}(I)$ of the six torus vibration modes

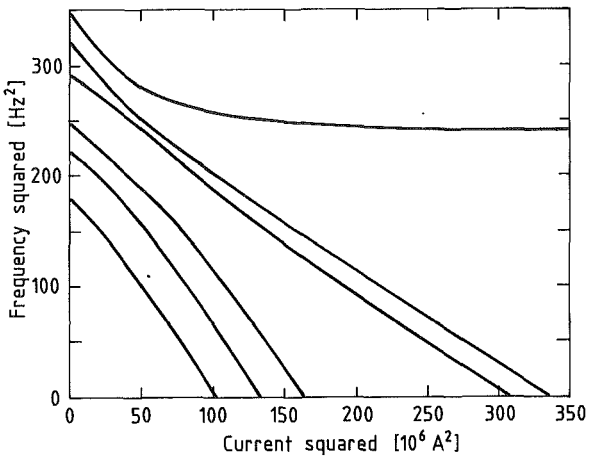


Fig. 7 Calculated frequency-current dispersion curves $f_{0i}^2(I^2)$ of the six torus vibration modes

(2) *Magnetic Stiffness Matrix.* As shown previously, the magnetic stiffness matrix can be calculated on the basis of the antiphase and in-phase differential inductances c_{q-} and c_{q+} at neighboring distances $q(q=0, 1, \ldots, N/2)$. For the basic deflection of TESPE, the in-phase differential mutual inductances are considerably smaller than the antiphase ones (by a factor of 40 for nearest neighbors) and therefore may be neglected in good approximation. The differential inductance c_{0+} is zero, as the self-inductance is not changed in the basic mode. Using equations (19) and (20), the magnetic stiffness matrix at a current I is then obtained as

$$\mathbf{K}^m = 2I^2 \begin{bmatrix} -\Sigma c_r & c_{1-} & c_{2-} & c_{3-} & c_{2-} & c_{1-} \\ c_{1-} & -\Sigma c_r & c_{1-} & c_{2-} & c_{3-} & c_{2-} \\ c_{2-} & c_{1-} & -\Sigma c_r & c_{1-} & c_{2-} & c_{3-} \\ c_{3-} & c_{2-} & c_{1-} & -\Sigma c_r & c_{1-} & c_{2-} \\ c_{2-} & c_{3-} & c_{2-} & c_{1-} & -\Sigma c_r & c_{1-} \\ c_{1-} & c_{2-} & c_{3-} & c_{2-} & c_{1-} & -\Sigma c_r \end{bmatrix} \tag{22}$$

where $\Sigma c_r := 2c_{1-} + 2c_{2-} + c_{3-}$. A magnetic field calculation program was used to determine the mutual inductances L_{rs} of the TESPE coils for varying deflection angles of basic deflection and different neighboring distances. By fitting the coefficients of the Taylor series (see equation (17)) were found. The following values were obtained for the antiphase differential inductances: $c_{1-} = 3.94 \times 10^{-3}$ H/rad², $c_{2-} = 4.80 \times 10^{-4}$ H/rad² and $c_{3-} = 2.09 \times 10^{-4}$ H/rad². It is obvious that the differential inductance between the coils and, hence, the magnetic stiffness, decreases considerably as the neighboring distance increases.

(3) *Frequency-Current Dispersion Curves and Buckling Current.* The frequencies of the torus vibration modes are obtained by solving the general matrix eigenvalue problem (cf. equation (16))

$$(-\omega_0^2 \mathbf{M} + (\mathbf{K}^e + \mathbf{K}^m))\hat{\phi}_0 = 0. \tag{23}$$

On the basis of the solutions for different current values I, the frequency-current dispersion curves of the six torus vibration modes are obtained (Fig. 6). The calculated frequencies of all torus modes decrease with increasing current, and the form of the torus modes (the relative vibration amplitudes of the coils) is also changed. The buckling current of the basic mode, i.e., the current at which the lowest natural frequency becomes zero, is calculated to be 10,080 A, and thus only 44 percent higher than the nominal operating current of 7000 A. Up to the operating current, the lowest frequency decreases from 13.4 Hz to 10.1 Hz, i.e., by 24 percent. A decrease of this magnitude can easily be verified experimentally.

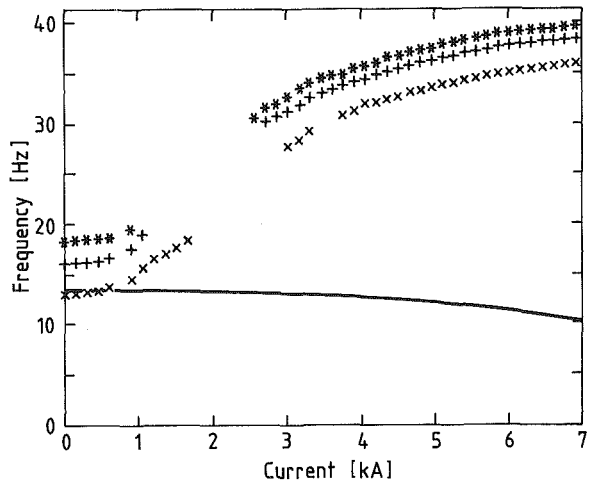


Fig. 8 Measured frequency-current dispersion curves of the torus modes (*, +, ×), contrasted with the calculated curve of the lowest torus mode (——)

In Fig. 7 the natural frequencies squared are plotted versus the TESPE current squared up to a current of 19,000 A. It is evident that the linear drop described in the literature (Moon, 1984; Miya and Uesaka, 1982) cannot be maintained when the differences in the elastic stiffnesses of the coils are taken into account. Therefore, the buckling current of magnet arrangements of this type cannot be determined from the slope of these curves at low currents as it was possible in earlier experimental investigations (Moon, 1979, Miya and Uesaka, 1982). If that determination method is applied to the curves in Fig. 7, a buckling current value of 12,350 A is obtained, which is too optimistic by 23 percent.

Overturning Mode. In case the TESPE coils are fixed by an azimuthal intercoil structure on the outside in the middle, the basic mode is excluded and the overturning mode becomes the mode with the lowest buckling current. The calculation of the buckling current which is significantly higher than the TESPE operating current is simplified for this mode here by assuming the elastic stiffnesses of the coils to be equal and by neglecting the elastic coupling between the coils. Then the torus mode with the lowest buckling current is the mode in which each coil deflects with the same amplitude but in the opposite direction as its nearest neighbors (Moon, 1984).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

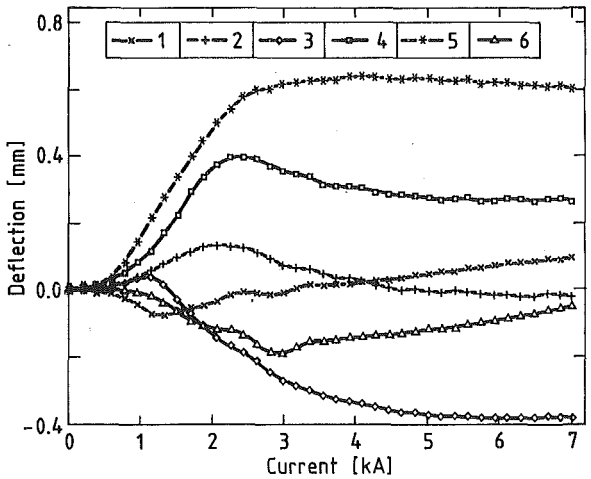


Fig. 9 Deflections of coils 1, 2, . . . , 6 at the outer part of the coils

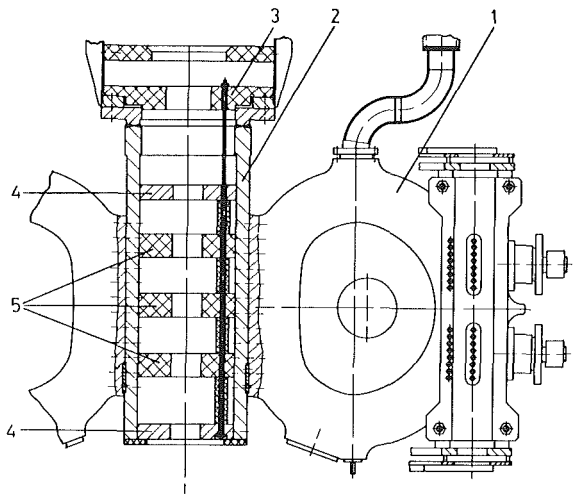


Fig. 10 Setup of the central support structure: 1—coil, 2—front part, 3—bearing ring, 4—top and bottom hexagonal disks, 5—inner disks

For the derivation of the buckling current, the deflection angles ϕ_i of the coils around their axis of symmetry are taken as coordinates of deflections and the torques acting on the coils are taken as generalized forces. In equation (12), the first term on the right side can be dropped, as it is independent of the deflection and, hence, has no influence on the buckling current. Using equation (12) with $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_6)$, and the vector of the magnetic torques on the coils being $\boldsymbol{\tau}^m = (\tau_1^m, \ldots, \tau_6^m)$, the equation

$$\boldsymbol{\tau}^m = -\mathbf{K}^m \boldsymbol{\phi} \tag{24}$$

is obtained. The magnetic stiffness matrix $\mathbf{K}^m$ was derived above (cf., equations (19), (20)). As far as the overturning mode is concerned, the differential self-inductance c_{0+} equals zero since the self-inductance is not changed by this deflection. In addition, the antiphase differential inductance c_{3-} of opposite coils is zero. Compared with the antiphase differential inductances c_{q-}, however, the in-phase differential mutual inductances c_{q+} are no longer negligibly small and have to be taken into account. Using equations (24), (19), and (20), the magnetic torque acting on the coils in case of the antiphase vibration then has the form

$$\tau_i^m = 8(c_{1-} + c_{2+}) I^2 \phi_i. \tag{25}$$

This magnetic torque has the same absolute value as the elastic torque $\tau_i^e = -K^e \phi_i$ when the current reaches the buckling value of

$$I_B = \sqrt{\frac{K^e}{8(c_{1-} + c_{2+})}}. \tag{26}$$

The differential inductances of the overturning mode were calculated numerically and amounted to $c_{1-} = 8.5 \times 10^{-4}$ H/rad^2 and $c_{2+} = -3.9 \times 10^{-4}$ H/rad^2. Using the natural frequency of the overturning vibration and the moment of inertia relative to the axis of symmetry, the elastic stiffness K^e was determined to be 6.1×10^7 Nm/rad. Using equation (26), a buckling current of 130 kA is obtained for the overturning mode. Compared with the TESPE operating current, this value is higher by a factor of 19. Compared with the buckling current calculated for the basic mode, it is higher by a factor of 13.

In his theoretical derivations for toroidal magnet systems, Moon always assumed that the forces between the coils only depend on the *differences* between the coil deflections. Consequently, the in-phase differential inductance of next nearest neighboring coils c_{2+} is neglected in his buckling current calculations (Moon, 1979, 1984). If this neglection would also be carried out for the overturning deflection of TESPE ($c_{2+} = 0$ in equation (26)), the resulting buckling current would be 95 kA or by 27 percent too low (see 130 kA above). It may be supposed that the deviations between the measured and calculated buckling currents of up to 39 percent observed by Moon (1979) in a 16-coil experiment result to a large fraction from the neglection of the in-phase differential inductances.

Measurements on TESPE and Discussion of Results

After having performed the precursor experiments using a special test coil and having achieved a very good agreement between experiment and theory, measurements of the buckling behavior of the TESPE magnet arrangement up to its operating current were carried out to show the justification of the above assumptions and numerical calculations. Moreover, it was the objective of the measurements to find out, whether other effects apart from magnetic stiffnesses contribute to a change of the lateral stiffnesses of the coils in a complex technical arrangement of this type. Especially, it was uncertain whether the foregoing assumption of a current-independent elastic stiffness was true in spite of the high central forces in a torus and the relatively complex construction of the TESPE central support structure.

For the measurements at low operating temperatures, the lateral deflections of the coils were detected by capacitive displacement transducers at the outer part of the coils. The coil vibrations were excited by the mechanical impact of a spring-loaded hammer within the cryostat, and the vibration modes were analyzed by applying the Fast Fourier Transform to the vibration signals.

Basic Mode. The basic vibrations were measured for a TESPE current increased stepwise to operating current. The frequencies of the identified torus modes are shown in Fig. 8. From the spectra of the vibration signals, three torus modes can be identified in the current range from zero to about 1500 A; damping of the vibrations increases with the current. Between 1500 A and 2500 A, damping of the vibrations is very large and the natural frequencies can no longer be determined on the basis of the measured signals. Above 2500 A, again three torus modes can be identified.

As a result of the calculations, a continuous decrease of the natural frequencies with increasing current was expected (cf., Fig. 6). However, the measurements showed a considerable *increase* in the frequencies after having exceeded about 800 A, and above 3500 A a flattening was observed. At 6900 A, the frequency of the lowest vibration mode is higher by a factor of 2.8 compared to that at zero current. The measured frequencies above 800 A clearly do not represent the reduction of the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

total stiffness expected from the buckling calculations and proved by the precursor experiments (cf., Fig. 4).

In Fig. 9 the measured static basic deflections of the six coils are plotted versus the current. On the basis of the calculated buckling current of 10,080 A and the very small misalignments of the coils at zero current, only small lateral deflections were expected at small currents (below 4000 A, not more than 0.1 mm in the outer part of the coils). Following the buckling theory, the deflections were supposed to increase with the current with a growing increment (cf., Fig. 3, BS coil). In fact, however, the coils already deflect significantly at currents above 800 A and instead of increasing with a growing increment, the curves flatten at higher currents. This unexpected behavior cannot be explained in terms of magnetoelastic buckling, either.

The deflections and stiffenings observed can be explained by the effect of the central magnetic forces in the torus, which press the coils onto the central support structure. In Fig. 10 the setup of the central support structure is shown. It was designed to have low mass to facilitate cooldown and to minimize eddy currents. Six flat steel pieces of 45 mm in thickness (the front parts) form a hollow hexagonal cylinder. This cylinder is reinforced by five hexagonal disks inside, the top and bottom of which are screwed to the front parts while the inner three are suspended from the bearing ring. The coils are screwed to the front parts. The complete torus is suspended from the top of the cryostat. At zero current, the front parts are somewhat bent to the outside as a result of the weight of the coils and the azimuthal intercoil structure. They are then only fixed to the top and bottom hexagonal disks (and to the bearing ring as part of the cryostat). With increasing current the torque around the bottom disk caused by the coil weight is more and more compensated by the torque resulting from the central magnetic force. At a current of 750 A, these two torques are equal and at higher currents the front parts are pressed against the inner three hexagonal disks. This additional support considerably increases the stiffness of torsional deflection, which leads to the observed frequency increase above about 800 A (cf., Fig. 8). At high currents, the "solidification" of the fixture and support structure and, hence, the frequency increase gradually comes to an end. The damping of the vibrations observed in the intermediate current range can be explained by the high friction at the inner three disks when the front parts are not yet strongly pressed against them.

The observed deflection behavior of the coils, too, can be explained by the effect of the central forces on the front parts (cf., Fig. 9). The front parts are not in tight contact with the inner hexagonal disks and not exactly parallel to them at zero current. With increasing current and forces the front parts are twisted and the small gaps between the front parts and inner disks are closed. This twist leads to a lateral deflection of the coils. The outlined explanation of the observed deflections and stiffenings was additionally confirmed by further measurements (Geiger, 1987).

Overturning Mode. The calculated buckling current of the overturning mode is so high that up to the operating current no measurable decrease in frequency as a result of negative magnetic stiffnesses can be expected. Therefore, the interest concentrated on the question whether the frequencies of the vibration modes would be changed by other effects. The measured frequencies continuously increase with the current. Between zero and 7000 A, the frequency of the lowest torus mode is increased by 26 percent and that of the highest mode by 21 percent. Obviously, the stiffening effect is considerably smaller in this mode compared with that in basic deflection.

Concluding Remarks

As required by the results of the performed investigations on magnetoelastic buckling, the hitherto existing models for the calculation of the magnetoelastic behavior of toroidal magnet systems were extended to obtain a more exact prediction of the buckling behavior of technical systems. Furthermore, it was shown that the elastic stiffnesses in a technical torus may vary considerably with the current, and that for this reason the buckling currents calculated with zero current elastic stiffnesses may be incorrect to a considerable extent. The buckling measurements also demonstrated the suitability of the natural frequencies of the torus system for monitoring its mechanical stability during operation.

Acknowledgments

This work has been performed in the framework of the Nuclear Fusion Project of the Kernforschungszentrum Karlsruhe and is supported by the European Communities within the European Fusion Technology Program. The authors thank H. Kiesel, G. Obermaier, and E. Süss for their skillful construction of various experimental components and for running the torus.

References

Chattopadhyay, S., 1979, "Magnetoelastic Instability of Structures Carrying Electric Current," *Int. J. Solids and Structures*, Vol. 15, pp. 467–477.

Geiger, W., 1987, "Magnetoelastisches Buckling in supraleitenden toroidalen Magnetsystemen," Kernforschungszentrum Karlsruhe KfK-4260, Dissertation Universität Karlsruhe.

Geiger, W., and Jüngst, K., P., 1987, "Magnetoelastic Buckling Measurements on the Superconducting Torus TESPE," *Electromagneto-mechanical Interactions in Deformable Solids and Structures, Proc. of the IUTAM Symp.*, Y. Yamamoto and K. Miya, eds., Elsevier Science Publishers, Amsterdam, pp. 21–26.

Jüngst, K. P., 1984, "Operation Experience with a 3-coil Arrangement of the TESPE Superconducting Torus," *Proc. of the 13th Symp. on Fusion Technology*, Pergamon Press, Oxford, pp. 1469–1474.

Jüngst, K. P., et al., 1985, "Superconducting Torus TESPE at Design Values," *Proc. of the 9th Int. Conf. on Magnet Technology*, C. Marinucci, ed., Swiss Inst. for Nuclear Research, Villingen, pp. 36–39.

Jüngst, K. P., Leppelmeier, G. W., Geiger, W., Komarek, P., and Steeves, M. M., 1986, "First Results of the TESPE-S Magnet System Safety Experiments," *Proc. of the 14th Symp. on Fusion Technology*, Pergamon Press, Oxford, Vol. 2, pp. 1759–1764.

Miya, K., Takagi, T., Uesaka, M., and Someya, K., 1982, "Finite Element Analysis of Magnetoelastic Buckling of Eight-Coil Superconducting Full Torus," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 180–186.

Miya, K., and Uesaka, M., 1982, "An Application of Finite Element Method to Magnetomechanics of Superconducting Magnets for Magnetic Fusion Reactors," *Nuclear Engineering and Design*, Vol. 72, pp. 275–296.

Moon, F. C., and Swanson, C., 1977, "Experiments on Buckling and Vibration of Superconducting Coils," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, pp. 707–713.

Moon, F. C., 1979, "Experiments on Magnetoelastic Buckling in a Superconducting Torus," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 145–150.

Moon, F. C., 1980, "Magnetoelastic Instabilities in Superconducting Structures and Earnshaw's Theorem," ASME *Journal of Mechanical Design*, Vol. 41, pp. 77–90.

Moon, F. C., 1984, *Magneto-Solid Mechanics*, John Wiley and Sons, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Usefulness of Elementary Theory for the Linear Vibrations of Layered, Orthotropic Elastic Beams and Corrections Due to Two-Dimensional End Effects[1]

J. M. Duva
Assoc. Mem. ASME

J. G. Simmonds
Mem. ASME

Department of Applied Mathematics,
University of Virginia,
Charlottesville, VA 22903

With the aid of formal asymptotic expansions, we conclude not only that elementary (Euler-Bernoulli) beam theory can be applied successfully to layered, orthotropic beams, possibly weak in shear, but also that, in computing the lower natural frequencies of a cantilevered beam, the most important correction to the elementary theory—of the relative order of magnitude of the ratio of depth to length—comes from effects in a neighborhood of the built-in end. We compute this correction using the fundamental work on semi-infinite elastic strips of Gregory and Gladwell (1982) and Gregory and Wan (1984). We also show that, except in unusual cases (e.g., a zero Poisson's ratio in a homogeneous, elastically isotropic beam), Timoshenko beam theory produces an erroneous correction to the frequencies of elementary theory of the relative order of magnitude of the square of the ratio of depth to length.

Introduction

Various authors have often asserted or implied that an adequate stress, vibration, or buckling analysis of a layered, orthotopic beam, especially one weak in shear, requires a refined or a higher-order beam theory. (See the review articles by Kapania and Raciti, 1989; Noor and Burton, 1989.) These arguments, which are frequently corroborated by numerical work, are based on the observation that, across the interfaces of the layers, the axial stress jumps and the displacements kink. Thus, elementary beam theory, being based on assuming simple, smooth thickness variations of the displacements, cannot predict the deviant behavior such discontinuities imply. However, our thesis is that such conclusions are based on a misunderstanding of the foundations and implications of elementary beam theory. Specifically, we claim (as have others before us) that the essence of a one-dimensional theory of the behavior of a two (or three-) dimensional continuum is that the variation of the unknowns along the relevant dimension is *slow* compared with their variation in the other (usually orthogonal) direction(s). Within the framework of linear plane elasticity, this implies, as we shall show, that in a narrow rectangular beam, a proper scaling of the unknowns leads to a set of partial differential equations that, to lowest order, may be reduced to the ordinary differential equation of elementary beam theory. Once this equation is solved, we may compute, with virtually no effort, a first approximation to the thickness variations of the displacements and stresses, even if there are kinks and jumps.

Further, we show (for elastically homogeneous isotropic beams only to make our point with a minimum of algebra) that the most important correction to the natural frequencies of vibration of a cantilevered beam comes from two-dimensional effects at the built-in end. This correction, which we compute using results of Gregory and Gladwell (1982) and Gregory and Wan (1984), is proportional to $\nu H/L$, where ν is Poisson's ratio, $2H$ is the depth of the beam, and L is its length. Finally, we show that, unless $\nu = 0$, the correction to the lower frequencies of vibration predicted by Timoshenko beam theory are erroneous, being proportional to H^2/L^2.

The Governing Equations

Consider a narrow rectangular beam which, if undeformed, occupies the region $0 \le x \le L$, $|y| \le b$, $|z| \le H$ in a fixed, inertial Cartesian reference frame, $Oxyz$. Let $[\sigma_x(x, z), \tau(x, z), \sigma_z(x, z)]$ denote the axial, shear, and normal stress parallel to the xz-plane, $[U(x, z), W(x, z)]$ axial and normal displacement, ρ the mass per unit volume of the beam, and t the time. Then,

[1]This work was supported in part by the National Science Foundation under Grant MSS-8903130 and in part by the Academic Enhancement Program at the University of Virginia.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 11, 1989; final revision, Feb. 5, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

in the linear theory of plane stress in the xz-plane, the equations of motion take the form

$$\sigma_{x,x}+\tau_{,z}=\rho U_{,tt} \tag{1}$$

$$\tau_{,x}+\sigma_{z,z}=\rho W_{,tt}, \tag{2}$$

where a comma followed by a subscript means differentiation with respect to that subscript.

The axial, shear, and normal strains are given in terms of the displacements by

$$\epsilon_x=U_{,x},\ \gamma=U_{,z}+W_{,x},\ \epsilon_z=W_{,z}. \tag{3}$$

To complete the field equations, we need constitutive relations. To make our point in as simple a way possible, we consider a beam made of an upper (+) layer and a lower (−) layer, each of constant thickness H with orthotropic strain-stress relations of the form

$$E_x^{\pm}\epsilon_x=\sigma_x-\nu_{\pm}\tilde{E}_{\pm}\sigma_z,\ G_{\pm}\gamma=\tau,\ E_z^{\pm}\epsilon_z=\sigma_z-\nu_{\pm}\sigma_x, \tag{4}$$

where $E_x^{\pm}$, $E_z^{\pm}$, $G_{\pm}$, and $\nu_{\pm}$ are the elastic moduli and Poisson ratios in the upper or lowers layers and

$$\tilde{E}=E_x/E_z. \tag{5}$$

Boundary Conditions. To be specific, we study the free vibrations of a cantilevered beam. Thus, the boundary conditions are

$$\textit{on the faces:}\qquad \tau(x,\pm H,t)=\sigma_z(x,\pm H,t)=0 \tag{6}$$

$$\textit{on the left end:}\qquad U(0,z,t)=W(0,z,t)=0 \tag{7}$$

$$\textit{on the right end:}\qquad \sigma_x(L,z,t)=\tau(L,z,t)=0. \tag{8}$$

Further, along the interface of the layers, we have the continuity conditions

$$[U]=[W]=[\tau]=[\sigma_z]=0, \tag{9}$$

where $[U]\equiv U(x, 0+, t) - U(x, 0-, t)$, etc.

Beam Theory

Elementary (Euler-Bernoulli) Theory. The equation of motion for the free vibrations of a cantilevered beam is

$$-BV_{,xxxx}=mV_{,tt} \tag{10}$$

and the boundary conditions are

$$V(0,t)=V_{,x}(0,t)=V_{,xx}(L,t)=V_{,xxx}(L,t)=0. \tag{11}$$

Here, B is the section modulus, V is some weighted measure of vertical deflection, and m times $2b$, the width of the beam, is the mass per unit length. The section modulus can be determined experimentally or by *descent from two-dimensions* if the composition of the beam is known. As we shall see, this latter procedure yields

$$B=\left[\frac{24E_x^+E_x^-+3(E_x^+-E_x^-)^2-2(E_x^++E_x^-)^2}{12(E_x^++E_x^-)}\right]H^3 \equiv(2/3)E_BH^3. \tag{12}$$

If $E_x^+ = E_x^- \equiv E$, then $E_B = E$.

Likewise, V may be related in various ways to the two-dimensional vertical displacement W. Here we follow Cowper (1966) and take V to be the *density-weighted average* of W through the thickness, i.e.,

$$V=\frac{\int_{-H}^{H}\rho W dz}{m}=\frac{\rho_-\int_{-H}^{0}Wdz+\rho_+\int_0^H Wdz}{(\rho_-+\rho_+)H}, \tag{13}$$

where $\rho_{\pm}$ is the constant density in the $\pm$-layer.

Timoshenko Theory. This comprises two *exact* equations of motion,

$$Q_{,x}=mV_{,tt},\ M_{,x}-Q=I\beta_{,tt}, \tag{14}$$

and two *approximate* constitutive relations,

$$M=B\beta_{,x},\ Q=kGH(V_{,x}+\beta). \tag{15}$$

In equations (14), which are derived from the equations of motion of plane stress theory by taking certain averages through the thickness, I is the moment of inertia per unit width of the beam,

$$Q=\int_{-H}^{H}\tau dz \text{ and } M=\int_{-H}^{H}z\sigma_x dz \tag{16}$$

are a shear force and moment *per unit width* of the beam, and

$$\beta=\frac{\int_{-H}^{H}z\rho U dz}{I}=\frac{\rho_-\int_{-H}^{0}zUdz+\rho_+\int_0^H zUdz}{(mH^2/3)} \tag{17}$$

is an angle of rotation (see, Cowper (1966) or Libai and Simmonds (1988)); in equation (15), k is a dimensionless shear coefficient and G is the shear modulus.

Note that the pointwise boundary conditions (7) and (8) imply that

$$V(0,t)=\beta(0,t)=M(L,t)=Q(L,t)=0. \tag{18}$$

Thus, the boundary conditions of Timoshenko theory are *exact* (but incomplete).

By differential elimination, we may reduce the field equations, (14) and (15), and the boundary conditions, (18), to

$$-BV_{,xxxx}=m[V+(mH/3kG)V_{,tt}-(H^2/3)(1+2E_B/kG)V_{,xx}]_{,tt}, \tag{19}$$

and

$$V(0,t)=0,\ V_{,t}(0,t)+(2/3)(E_BH^2/kG)[V_{,xxx}(0,t)-(m/kGH)V_{,xtt}(0,t)]=0 \tag{20}$$

$$V_{,xx}(L,t)-(m/kGH)V_{,tt}(L,t)=0,$$

$$V_{,xxx}(L,t)-[(m/kGH)+(m/2E_BH)]V_{,xtt}(L,t)=0. \tag{21}$$

Later, we shall compare the natural frequencies of vibration predicted by Timoshenko theory with those of elementary theory and with those of elementary theory corrected for end effects.

Nondimensionalization for Free Vibrations

We now assume that the beam executes free harmonic motion at frequency Ω, "amplitude" W_0, and "wavelength" l. Beam theory applies so long as H/l is small. However, as l is not known *a priori* (as it would be if the beam were simply supported or infinite), we shall take, instead,

$$\epsilon=H/L \tag{22}$$

as our basic small parameter. Then, beam theory holds if $\epsilon \ll 1$ and $L/l = O(1)$ and we may expect the same relative orders of magnitude of the axial, shear, and normal stress as in static theory. These observations and the structure of the equations of motion, (1) and (2), the strain-displacement relations, (3), and the strain-stress relations, (4), suggest that we introduce dimensionless variables and a dimensionless frequency by setting

$$x=L\xi,\ z=H\zeta \tag{23}$$

$$(U,W)=W_0(\epsilon u,w)e^{i\Omega t},\ (\sigma_x,\tau,\sigma_z)=(E_BH/L^2)W_0(s_\xi,\epsilon s,\epsilon^2 s_\zeta)e^{i\Omega t} \tag{24}$$

$$\Omega^2=(2E_BH^3/3mL^4)\omega^2. \tag{25}$$

Then, rearranging (1)–(4) in the form and order in which we shall integrate them, with strains expressed in terms of displacements and with the dimensionless axial stress,

$$s_\xi=(E_x^{\pm}/E_B)u_{,\xi}+(\nu\tilde{E})_{\pm}\epsilon^2 s_\zeta \tag{26}$$

eliminated, we obtain the field equations

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$w_{,\zeta} = -(\nu E_x/E_z)_{\pm}\epsilon^2 u_{,\xi} + (E_B/E_z^{\pm})\epsilon^4[1-(\nu^2\tilde{E})_{\pm}]s_{\zeta} \quad (27)$$

$$u_{,\zeta} = -w_{,\xi} + (E_B/G_{\pm})\epsilon^2 s \quad (28)$$

$$s_{,\zeta} = -(E_x^{\pm}/E_B)u_{,\xi\xi} - \epsilon^2[(2\rho_{\pm}H\omega^2/3m)u + (\nu\tilde{E})_{\pm}s_{\zeta,\xi}] \quad (29)$$

$$s_{\zeta,\zeta} = -s_{,\xi} - (2\rho_{\pm}H/3m)\omega^2 w. \quad (30)$$

The traction-free conditions on the faces of the beam, equation (6), and the continuity conditions at the interface, equation (9), reduce to

$$s(\xi,\pm 1) = s_{\zeta}(\xi,\pm 1) = 0 \quad (31)$$

$$[u] = [w] = [s] = [s_{\zeta}] = 0, \quad (32)$$

while the end conditions, equations (7) and (8), become

$$u(0,\zeta) = w(0,\zeta) = 0 \quad (33)$$

$$(E_x^{\pm}/E_B)u_{,\xi}(1,\zeta) + (\nu\tilde{E})_{\pm}\epsilon^2 s_{\zeta}(1,\zeta) = s(1,\zeta) = 0. \quad (34)$$

Interior Asymptotic Solution for Comparable Extensional and Shear Moduli

The field equations, (27)–(30), and the boundary and jump conditions, (31)–(34), constitute an eigenvalue problem. Their structure suggests that, as long as derivatives with respect to ξ and ζ are O(1), i.e., in the *interior* of the beam away from ends, we seek solutions for each unknown of the form

$$f^I = f^{(0)}(\xi,\zeta) + \epsilon^2 f^{(1)}(\xi,\zeta) + \ldots . \quad (35)$$

(We shall drop the superscript I in this section.) Likewise, we seek an expansion for the squares of the dimensionless natural frequencies of the form

$$\omega^2 = \omega_0^2 + \epsilon^2\omega_1^2 + \ldots . \quad (36)$$

As we shall see, these are *not* the correct expansions (which proceed in powers of ϵ and not ϵ^2), even though only ϵ^2 appears in our eigenvalue problem. Nevertheless, to motivate the correct expansion, we shall adopt (35) and (36) until trouble arises. (Straying down the wrong path for a while will cost us little for we shall use most of what we compute.)

Lowest-Order Approximations. To lowest order, equation (27) implies that $w^{(0)}{}_{,\zeta} = 0$; hence,

$$w^{(0)} = \nu_0(\xi), \quad (37)$$

where the unknown function ν_0 must be continuous at $\zeta = 0$ because $[w] = 0$. From $(33)_2$, $\nu_0(0) = 0$.

To lowest order, equation (28) now yields

$$u^{(0)} = -\nu_0'(\xi)\zeta + u_0^{\pm}(\xi), \quad (38)$$

where $u_0^{\pm}$ are unknown functions of integration. However, because $[u] = 0$, $u_0^+ = u_0^- \equiv u_0$. Next, we insert equation (38) into (29) and integrate with respect to ζ to obtain, to lowest order,

$$s^{(0)} = C_{\pm}[\nu_0'''(\xi)(\zeta^2/2) - u_0''(\xi)\zeta] + s_0^{\pm}(\xi), \quad (39)$$

where

$$C_{\pm} = E_x^{\pm}/E_B \quad (40)$$

and $s_0^{\pm}$ are functions of integration. Because $[s] = 0$, $s_0^+ = s_0^- \equiv s_0$. The vanishing of the dimensionless shear stress on the faces then yields

$$u_0 = D\nu_0'(\xi) + A_0\xi + B_0,\ s_0 = -P\nu_0'''(\xi), \quad (41)$$

where

$$D = \frac{1}{2}\left(\frac{C_+ - C_-}{C_+ + C_-}\right),\ P = \frac{C_+C_-}{C_+ + C_-}, \quad (42)$$

and A_0 and B_0 are unknown constants.

The vanishing of the axial displacement on the left end of the beam implies that $u^{(0)}(0,\zeta) = 0$ and hence, by equations (38) and (41), that $B_0 = \nu_0'(0) = 0$; moreover, if we substitute equations (38) and (39) into (34), the stress-free conditions at the right end of the beam, we find that $A_0 = \nu_0''(1) = \nu_0'''(1) = 0$. Thus,

$$u^{(0)} = \nu_0'(\xi)(D-\zeta) \quad (43)$$

and

$$\begin{aligned} s^{(0)} &= -\nu_0'''(\xi)[P + C_{\pm}\zeta(D-\zeta/2)] \\ &= -\nu_0'''(\xi)(1\mp\zeta)(P \pm C_{\pm}\zeta/2) \\ &= -(1/2)\nu_0'''(\xi)(1-\zeta^2) \text{ if } E_x^+ = E_x^- = E. \end{aligned} \quad (44)$$

Finally, we turn to (30), insert (36), (37), and (44), and integrate with respect to ζ. To lowest order, this yields

$$\begin{aligned} s_{\zeta}^{(0)} = \nu_0'''(\xi)[P + (1/2)C_{\pm}\zeta(D-\zeta/3)]\zeta \\ - (2\rho_{\pm}H/3m)\omega_0^2\nu_0(\xi)\zeta + s_{\zeta 0}^{\pm}(\xi), \end{aligned} \quad (45)$$

where $s_{\zeta 0}^{\pm}$ are unknown functions of integration. Because $[s_{\zeta}] = 0$, $s_{\zeta 0}^+ = s_{\zeta 0}^- \equiv s_{\zeta 0}$. Imposing the face boundary conditions, $s_{\zeta}(\xi, \pm 1) = 0$, and noting equation (12), we find that ν_0 and ω_0 satisfy the *dimensionless elementary beam equation*,

$$-\nu_0'''' + \omega_0^2\nu_0 = 0, \quad (46)$$

so that

$$\begin{aligned} s_{\zeta}^{(0)} &= \omega_0^2\nu_0(\xi)\left[\frac{C_- - C_+}{24} + \frac{(\rho_+ - \rho_-)H}{3m} + \left(P - \frac{2\rho_{\pm}H}{3m}\right)\zeta \right. \\ &\left. + \frac{C_{\pm}}{2}\zeta^2(D-\zeta/3)\right] = \omega_0^2\nu_0(\xi)(1\mp\zeta)\left[\frac{C_- - C_+}{24} + \frac{(\rho_+ - \rho_-)H}{3m} \right. \\ &\left. + \left(P - \frac{1}{3} \pm \frac{C_- - C_+}{24}\right)\zeta \pm \frac{C_{\pm}\zeta^2}{6}\right] \\ &\equiv (1/6)\omega_0^2\nu_0(\xi)(\zeta - \zeta^3) \text{ if } E_x^{\pm} = E_x^- = E. \end{aligned} \quad (47)$$

The emergence of the boundary conditions

$$\nu_0(0) = \nu_0'(0) = \nu_0''(1) = \nu_0'''(1) = 0 \quad (48)$$

has already been noted. Mode shapes (eigenfunctions) and the associated values of the dimensionless natural frequencies (square roots of the eigenvalues) are given by Blevins (1979).

Next-Order Approximation for Homogeneous Isotropic Beams. To avoid cluttering our approach with algebraic details, we restrict attention in this subsection to an elastically homogeneous isotropic beam. (We shall mention the novel problem of beams weak in shear near the end of the paper.) Thus, $D = 0$ and $P = \frac{1}{2}$, so that, substituting equation (43) into the right side of equation (27) and integrating with respect to ζ, we obtain

$$w^{(1)} = (\nu/2)\nu_0''(\xi)\zeta^2 + \nu_1(\xi), \quad (49)$$

where ν_1 is an unknown function of integration. The price of neglecting end effects is that there is no way we can choose $\nu_1(0)$ to make $w^{(1)}(0,\zeta) = 0$, unless $\nu = 0$ or, by accident, $\nu_0''(0) = 0$.

Next, we substitute equations (44) and (49) into (28) and integrate with respect to ζ to obtain

$$u^{(1)} = -\nu_0'''(\xi)[(1+\nu)\zeta - (2+\nu)\zeta^3/6] - \nu_1'(\xi)\zeta, \quad (50)$$

there being no function of integration because u — and hence $u^{(1)}, u^{(2)}$, etc.—must be *odd* in ζ in an isotropic beam undergoing pure transverse vibrations. Note that, except in unusual circumstances, there is no way we can choose $\nu_1'(0)$ so that $u^{(1)}(0,\zeta) = 0$. Furthermore, by equations (26), (46), (47), and (50),

$$s_{\xi}^{(1)} = -[(1/6)\omega_0^2\nu_0(\xi)(6 + 5\nu - 2\zeta^2) + \nu_1''(\xi)]\zeta. \quad (51)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Thus, there is no way that we can choose ν_1'' to make $s_\xi^{(1)}$ vanish on the right end of the beam.

To obtain $s^{(1)}$, we substitute equations (36), (43), (47), and (50) into (29), integrate with respect to ζ, and note equation (46). As there must be no shear stress on the faces of the beam, we find that

$$s^{(1)} = -(1/2)[(1/6)\omega_0^2 \nu_0'(\xi)(8+5\nu-\zeta^2) + \nu_1'''(\xi)](1-\zeta^2). \quad (52)$$

We cannot, in general, choose ν_1''' (1) so that $s^{(1)}$ vanishes at the right end of the beam.

End Effects

It is senseless to proceed without discussing end effects. For the relatively low frequencies of vibration we are considering, these effects should be confined to end zones of width $O(H)$, as suggested by the useful discussions by Boley (1955, 1960) and Grandin and Little (1974) of a dynamic St-Venant's Principle for a semi-infinite elastic strip.

Thus, we introduce the left and right edge-zone coordinates

$$\alpha = \frac{x}{H} = \frac{\xi}{\epsilon} \text{ and } \beta = \frac{L-x}{H} = \frac{1-\xi}{\epsilon}. \quad (53)$$

An order-of-magnitude analysis then suggests that we seek solutions, in terms of the dimensionless variables defined equations (23) and (24), of the form

$$U = W_0[\epsilon u^I(\xi, \zeta, \epsilon) + \epsilon^2 u^L(\alpha, \zeta, \epsilon) - \epsilon^4 u^R(\beta, \zeta, \epsilon)]e^{i\Omega t} \quad (54)$$

$$W = W_0[w^I(\xi, \zeta, \epsilon) + \epsilon^2 w^L(\alpha, \zeta, \epsilon) + \epsilon^4 w^R(\beta, \zeta, \epsilon)]e^{i\Omega t} \quad (55)$$

$$\sigma_x = (E_B H W_0/L^2)[s_\xi^I(\xi, \zeta, \epsilon) + s_\xi^L(\alpha, \zeta, \epsilon) + \epsilon^2 s_\xi^R(\beta, \zeta, \epsilon)]e^{i\Omega t} \quad (56)$$

$$\tau = (E_B H W_0/L^2)[\epsilon s^I(\xi, \zeta, \epsilon) + s^L(\alpha, \zeta, \epsilon) - \epsilon^2 s^R(\beta, \zeta, \epsilon)]e^{i\Omega t} \quad (57)$$

$$\sigma_z = (E_B H W_0/L^2)[\epsilon^2 s_\zeta^I(\xi,\zeta,\epsilon) + s_\zeta^L(\alpha,\zeta,\epsilon) + \epsilon^2 s_\zeta^R(\beta,\zeta,\epsilon)]e^{i\Omega t}. \quad (58)$$

In these equations, the superscripts L and R stand for "left" and "right" and we assume that all left and right edge-zone solutions decay as α or β approach infinity. The order of magnitude of the unknowns in the left and right end-zones is dictated by the fact that, to lowest order, the interior variables satisfy all boundary conditions and by the (erroneous) assumption that the interior unknowns have asymptotic expansions in powers of ϵ.[2]

The Left End-Zone Problem. Assuming that $\omega = O(1)$, invoking linearity, and noting the decomposition (54)–(58), we find that the equations of motion, (1) and (2), and the stress-strain relations, (4), take the form

$$s_{\xi,\alpha}^L + s_{,\zeta}^L = O(\epsilon^4) \quad (59)$$

$$s_{,\alpha}^L + s_{\zeta,\zeta}^L = O(\epsilon^4) \quad (60)$$

$$u_{,\alpha}^L = s_\xi^L - \nu s_\zeta^L,\ w_{,\alpha}^L + u_{,\zeta}^L = 2(1+\nu)s^L,\ w_{,\zeta}^L = s_\zeta^L - \nu s_\xi^L. \quad (61)$$

The face boundary conditions read

$$s^L(\alpha, \pm 1, \epsilon) = s_\zeta^L(\alpha, \pm 1, \epsilon) = 0 \quad (62)$$

and, because $\nu_0(0) = \nu_0'(0) = 0$ and from equations (35), (49), (50), (54), and (55), the conditions of no displacement on the left end for the beam take the form

$$w^L(0,\zeta,\epsilon) = -[\nu_1(0) + (\nu/2)\nu_0''(0)\zeta^2] + O(\epsilon^2) \equiv \bar{w}(\zeta,\epsilon) \quad (63)$$

$$u^L(0,\zeta,\epsilon) = \epsilon\{\nu_1'(0)\zeta + \nu_0'''(0)[(1+\nu)\zeta - (2+\nu)\zeta^3/6]\} + O(\epsilon^3) \equiv \bar{u}(\zeta,\epsilon). \quad (64)$$

[2]The correct representations turn out to be equations (70)–(74), but this is not yet apparent.

To lowest order, (59)–(64) are the *static* equations for an end-loaded, elastically isotropic strip. In a profound study of end effects in such strips, Gregory and Wan (1984) have shown that a necessary and sufficient condition for the prescribed end displacements $\hat{u}$ and $\hat{w}$ to produce an *exponentially decaying* state of stress is that

$$\int_{-1}^{1}[\sigma_x^{T,B,F}(\zeta)\bar{u}(\zeta,\epsilon) + \tau^{T,B,F}(\zeta)\bar{w}(\zeta,\epsilon)]d\zeta = 0, \quad (65T, B, F)$$

where $\sigma_x^{T,B,F}$ and $\tau^{T,B,F}$ are certain dimensionless *canonical* end stresses that arise in a semi-infinite, elastically isotropic strip in a state of unit tension (T), bending (B), or flexure (F) at infinity. We note that these canonical problems are normalized so that

$$\int_{-1}^{1}\sigma_x^T(\zeta)d\zeta = \int_{-1}^{1}\zeta\sigma_x^B(\zeta)d\zeta = \int_{-1}^{1}\tau^F(\zeta)d\zeta = 1. \quad (66)$$

Furthermore,

$$\int_{-1}^{1}\zeta\sigma_x^{T,F}(\zeta)d\zeta = \int_{-1}^{1}\tau^{T,B}(\zeta)d\zeta = 0 \quad (67)$$

and

$$\tau^T(\zeta) = -\tau^T(-\zeta) \text{ and } \tau^B(\zeta) = \tau^B(-\zeta). \quad (68)$$

The numerical computation of $\sigma_x^{T,B,F}(\zeta)$ and $\tau^{T,B,F}(\zeta)$ is discussed by Gregory and Gladwell (1982).

Inserting (63) and (64) into (65B) and noting (67), we find, to lowest order (i.e., as $\epsilon \to 0$) that

$$(\nu/2)\nu_0''(0)\int_{-1}^{1}\zeta^2\tau^B(\zeta)d\zeta = 0. \quad (69)$$

This is a contradiction, unless $\nu = 0$ or, by accident, $\nu_0''(0) = 0$.

The Correct Expansions

A little thought suggests that the contradiction in equation (69) can be avoided if, instead of (54)–(58), we assume that the displacements and stresses have the representations

$$U = W_0[\epsilon u^I(\xi,\zeta,\epsilon) + \epsilon u^L(\alpha,\zeta,\epsilon) - \epsilon^3 u^R(\beta,\zeta,\epsilon)]e^{i\Omega t} \quad (70)$$

$$W = W_0[w^I(\xi,\zeta,\epsilon) + \epsilon w^L(\alpha,\zeta,\epsilon) + \epsilon^3 w^R(\beta,\zeta,\epsilon)]e^{i\Omega t} \quad (71)$$

$$\sigma_x = (E_B H W_0/L^2)[s_\xi^I(\xi,\zeta,\epsilon) + s_\xi^L(\alpha,\zeta,\epsilon) + \epsilon s_\xi^R(\beta,\zeta,\epsilon)]e^{i\Omega t} \quad (72)$$

$$\tau = (E_B H W_0/L^2)[\epsilon s^I(\xi,\zeta,\epsilon) + s^L(\alpha,\zeta,\epsilon) - \epsilon s^R(\beta,\zeta,\epsilon)]e^{i\Omega t} \quad (73)$$

$$\sigma_z = (E_B H W_0/L^2)[\epsilon^2 s_\zeta^I(\xi,\zeta,\epsilon) + s_\zeta^L(\alpha,\zeta,\epsilon) + \epsilon s_\zeta^R(\beta,\zeta,\epsilon)]e^{i\Omega t}. \quad (74)$$

The New Interior Expansions. In place of equation (35), we now seek expansions for each interior unknown of the form

$$f^I = f^{(0)}(\xi,\zeta) + \epsilon f^{(1)}(\xi,\zeta) + \ldots. \quad (75)$$

Likewise, we assume that the square of the dimensionless frequency can be expressed as

$$\omega^2 = \omega_0^2 + \epsilon\omega_1^2 + \ldots. \quad (76)$$

It may be verified that if we insert such expansions into the dimensionless field equations (27)–(30) and specialize to an isotropic beam, then we obtain, as before, the same equations to lowest order. However, in place of equations (49)–(52), we obtain,

$$w^{(1)} = \nu_1(\xi),\ u^{(1)} = -\nu_1'(\xi)\zeta \quad (77)$$

$$s_\xi^{(1)} = -\nu_1''(\xi)\zeta,\ s^{(1)} = -(1/2)\nu_1'''(\xi)(1-\zeta^2). \quad (78)$$

So far, the first-order terms have the same from as the lower-order ones. However, equation (30) yields

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$s_\zeta^{(1)} = (1/6)\nu_1''''(\xi)(3-\zeta^2)\zeta - (1/3)[\omega_0^2\nu_1(\xi) + \omega_1^2\nu_0(\xi)]\zeta. \quad (79)$$

Satisfaction of the face boundary condition, $s_\zeta(\xi, \pm 1, \epsilon) = 0$, implies that

$$\nu_1'''' - \omega_0^2\nu_1 = \omega_1^2\nu_0(\xi) \quad (80)$$

and hence

$$s_\zeta^{(1)} = (1/6)[\omega_0^2\nu_1(\xi) + \omega_1^2\nu_0(\xi)](1-\zeta^2)\zeta. \quad (81)$$

Let λ_i denote the square of the ith natural frequency of vibration predicted by elementary beam theory—see equations (46) and (48)—and let $\phi_i(\xi)$ denote the associated natural mode, normalized so that

$$\int_0^1 \phi_i(\xi)\phi_j(\xi)d\xi = \delta_{ij}, \quad (82)$$

where δ_{ij} is the Kronecker delta. As these natural modes form a complete set, the solution of equation (80) can be represented as

$$\nu_1 = \sum_0^\infty a_n\phi_n(\xi). \quad (83)$$

For equation (80) to have a solution if $\omega_0^2 = \lambda_i$, ω_1 must be chosen so that

$$\int_0^1 [\nu_1''''(\xi) - \lambda_i\nu_1(\xi)]\phi_i(\xi)d\xi = \omega_1^2. \quad (84)$$

On integrating the first term on the left four times by parts and evoking equation (80) and the boundary conditions on ϕ_i, namely equation (48), we find that

$$\omega_1^2 = \nu_1''''(1)\phi_i(1) - \nu_1''(1)\phi_i'(1) - \nu_1'(0)\phi_i''(0) + \nu_1(0)\phi_i'''(0). \quad (85)$$

This choice of ω_1^2 leaves the coefficient a_i in equation (83) arbitrary. However, if we require that

$$\int_0^1 \nu^2(\xi)d\xi = \int_0^1 [\nu_0(\xi) + \epsilon^2\nu_1(\xi) + \ldots]^2 d\xi = 1, \quad (86)$$

then $a_i = 0$, and the other a_n's may be determined by standard procedures.

Finally, going to second order, we find that $w^{(2)}$ and $u^{(2)}$ are given, essentially, by equations (49) and (50), except that, now,

$$w^{(2)} = (\nu/2)\nu_0''(\xi)\zeta^2 + \nu_2(\xi) \quad (87)$$

$$u^{(2)} = -\nu_0'''(\xi)[(1+\nu)\zeta - (2+\nu)\zeta^3/6] - \nu_2'(\xi)\zeta. \quad (88)$$

To determine boundary conditions on ν_1 (and hence ω_1^2), we consider:

Edge Effects Anew

Because $w^{(0)}(0, \zeta) = \nu_0(0) = 0$ and $u^{(0)}(0, \zeta) = -\nu_0'(0)\zeta = 0$, it follows from equations (70), (71), and (77) that the conditions of zero displacement on the left end of the beam take the form

$$w^L(0, \zeta, \epsilon) = -\nu_1(0) - \epsilon[(\nu/2)\nu_0''(0)\zeta^2 + \nu_2(0)] + O(\epsilon^2) \equiv \bar{w}(\zeta) \quad (89)$$

$$u^L(0, \zeta, \epsilon) = \epsilon\nu_1'(0)\zeta + O(\epsilon^2) \equiv \bar{u}(\zeta). \quad (90)$$

Substituting equations (89) and (90) into (65) and noting (66)–(68), we find that, *to lowest order*, (65*T*) and (65*B*) are satisfied identically while (65*F*) reduces to

$$\nu_1(0) = 0. \quad (91)$$

To first order, equation (65*T*) is satisfied identically while (65*B*) yields

$$\nu_1'(0) = (\nu/2)\nu_0''(0)\int_{-1}^1 \zeta^2\tau^B(\zeta)d\zeta \equiv \nu C_B(\nu)\nu_0''(0); \quad (92)$$

equation (65*F*) yields an analogous equation for $\nu_2(0)$.

To get values for $\nu_1''(1)$ and $\nu_1'''(1)$ we consider the right end zone. To lowest order, the dimensionless field equations and the stress-free face conditions are the same as in the left end zone, equations (59)–(62), save that the superscript L is replaced by R.

Setting $s_\xi^{(0)}(1, \zeta) = s^{(0)}(1, \zeta) = 0$ gave us the boundary conditions $\nu_0''(1) = \nu_0'''(1) = 0$. Hence, by equations (72), (73), and (78), the stress-free conditions at the right end of the beam reduce to

$$s_\xi^R(0, \zeta, \epsilon) = \nu_1''(1)\zeta + O(\epsilon),\ s^R(0, \zeta, \epsilon) = -(\epsilon/2)\nu_1'''(1)(1-\zeta^2) + O(\epsilon^2). \quad (93)$$

But if the right end-zone solutions are to decay as $\beta \to \infty$, the stresses on the left sides of equations (93) must be self-equilibrated; hence

$$\nu_1''(1) = \nu_1'''(1) = 0. \quad (94)$$

It now follows from equations (85), (91), (92), and (94) that

$$\omega_1^2 = -\nu C_B(\nu)[\phi_i''(0)]^2.$$

From Table 8-2(b) of Blevins (1979), $\phi_i''(0) = 2\omega_0$. Thus,

$$(\omega/\omega_0)^2 = 1 - 4\nu C_B(\nu)\epsilon + O(\epsilon^2). \quad (96)$$

From the first and second columns of Table 4 of Gregory and Gladwell (1982), $C_B(1/4) = 0.07948$, $C_B(1/3) = 0.11361$, and $C_B(1/2) = 0.20351$.

The Corrections to Elementary Theory Predicted by Timoshenko Theory

Setting $x = L\xi$,

$$V = W_0\nu(\xi)e^{i\Omega t}, \quad (97)$$

and specializing to an elastically homogeneous isotropic beam so that equation (25) reads

$$\Omega^2 = (EH^2/3\rho L^4)\omega^2, \quad (98)$$

we find that the Timoshenko beam equation (19) reduces to

$$\nu'''' - \omega^2\nu + (1/3+q)\epsilon^2\omega^2\nu'' + (q/3)\epsilon^4\omega^4\nu = 0, \quad (99)$$

while the boundary conditions (20) and (21) become

$$\nu(0) = 0,\ \nu'(0) + q\epsilon^2[\nu'''(0) + q\epsilon^2\nu'(0)] = 0 \quad (100)$$

$$\nu''(1) + q\epsilon^2\omega^2\nu(1) = 0,\ \nu'''(1) + (1/3+q)\epsilon^2\omega^2\nu'(1) = 0, \quad (101)$$

where

$$q = \frac{4(1+\nu)}{3k}. \quad (102)$$

(Recall that k is the shear coefficient.)

The form of these equations suggest that we expand in powers of ϵ^2, as in equations (35) and (36). Doing so, we find that (ν_0, ω_0) satisfies the eigenvalue problem of elementary beam theory, equations (46) and (48), while (ν_1, ω_1) satisfies the differential equation

$$\nu_1'''' - \omega_0^2\nu_1 = \omega_1^2\nu_0 - (1/3+q)\omega_0^2\nu_0'' \quad (103)$$

and boundary conditions

$$\nu_1(0) = 0,\ \nu_1'(0) = -q\nu_0'''(0) \quad (104)$$

$$\nu_1''(1) = -q\omega_0^2\nu_0(1),\ \nu_0'(1) = -(1/3+q)\omega_0^2\nu_0'(1). \quad (105)$$

Note that if we take for the shear factor $k = 5/3$, then, save for a Poisson ratio term, we have the same boundary value problem that would have emerged from our original asymptotic calculation of $w^{(1)}, \ldots, s_\zeta^{(1)}$ (based on the incorrect expansions (35) and (36)) had we imposed the integrated end conditions $\int_{-1}^1 w^{(1)}(0, \zeta)d\zeta = 0$ and $\int_{-1}^1 \zeta u^{(1)}(0, \zeta)d\zeta = 0$ instead of the Gregory-Wan decay conditions (65). Thus, Timoshenko theory gives an $O(\epsilon^2)$ correction to the dimensionless frequencies of elementary theory. This is not an artifact of the arbitrariness of the shear factor, k, but rather a characteristic of

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

any higher-order or "refined" theory that fails to account for the clamped condition properly.

As a referee has pointed out, assuming expansions in powers of ϵ rather than ϵ^2 will not remedy the above defect in the Timoshenko beam equations; such an expansion only yields zeros for the coefficients of the ϵ-term.

End Effects in Beams Weak in Shear

Suppose that $G/E = \epsilon A$, $A = O(1)$. If we introduce a dimensionless Airy-stress-function, f, then, in the left end zone

$$f_{,\alpha\alpha\alpha\alpha} + (\epsilon^{-1}A^{-1} - 2\nu)f_{,\alpha\alpha\zeta\zeta} + f_{,\zeta\zeta\zeta\zeta} = O(\epsilon^4), \qquad (106)$$

the terms on the right being inertial. To lowest order, we have the *hyperbolic* equation $f^{(0)}_{,\alpha\alpha\zeta\zeta} = 0$, which does not exhibit decaying solutions. To avoid this, we introduce a scaled axial coordinate

$$\gamma = \sqrt{\epsilon A}\,\alpha \qquad (107)$$

which leads, to lowest order, to the Poisson equation

$$f^{(0)}_{,\gamma\gamma} + f^{(0)}_{,\zeta\zeta} = P(\gamma) + Q(\gamma)\zeta, \qquad (108)$$

where P and Q are unknown functions.

As pointed out by Choi and Horgan (1977) and Horgan (1982), end effects now penetrate much further into the interior, a dimensional distance of $O(H\sqrt{E/G})$. As a consequence, ϵ must be replaced everywhere by $\epsilon^{1/2}$ in our representations (70)–(74) and the interior variables and ω^2 have expansions in powers of $\epsilon^{1/2}$. Thus, for beams weak in shear in the sense that $G/E = O(H/L)$, effects at the clamped end may produce a relative error in the natural frequencies predicted by elementary theory of $O(\sqrt{H/L})$. We leave the quantification of this correction for a future study.

Conclusions

Our results cast certain doubts on the significance and practical value of many papers that develop refined theories of laminated beams, plates, or shells with the aim of removing deficiencies in classical theories. Often, the claims of improved accuracy are based on comparisons with two- or three-dimensional elasticity theory for *simply-supported* bodies. However, as boundary conditions of simple support for finite bodies are equivalent to symmetry conditions at nodes (or antinodes) of infinite bodies under periodic loads, such comparisons do not exhibit edge effects. As we have shown for the special (but suggestive) case of a cantilevered beam, edge effects are the source of the biggest errors in classical theories, and the weaker the beam in shear the more pronounced these errors. At the same time, we caution that our ϵ-end-effect corrections to elementary beam theory are *asymptotic* corrections, valid only as $\epsilon = H/L \to 0$. For small but finite values of ϵ, it may well be that the coefficients of the ϵ^2-terms are much larger than those of the ϵ-terms. Indeed, this is what Gregory and Wan (1984) found for a built-in, semi-infinite strip under pressure on its long sides.

Acknowledgments

We thank our colleague, C. O. Horgan, for a careful reading and suggestions for improvements of the manuscript.

References

Blevins, R. D., 1979, *Formulas for Natural Frequency and Mode Shape*, van Nostrand Reinhold, New York.

Boley, B. A., 1955, "Application of Saint-Venant's Principle in Dynamical Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, pp. 204–206.

Boley, B. A., 1960, "On a Dynamical Saint-Venant's Principle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 74–78.

Choi, I., and Horgan, C. O., 1977, "Saint-Venant's Principle and End Effects in Anisotropic Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, pp. 424–430.

Cowper, G. R., 1966, "The Shear Coefficient in Timoshenko's Beam Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 335–340.

Grandin, H. T., and Little, R. W., 1974, "Dynamic Saint-Venant Region in a Semi-Infinite Elastic Strip," *J. Elasticity*, Vol. 4, pp. 131–146.

Gregory, R. D., and Gladwell, I., 1982, "The Cantilever Beam Under Tension, Bending, or Flexture at Infinity," *J. Elasticity*, Vol. 12, pp. 317–343.

Gregory, R. D., and Wan, F. Y. M., 1984, "Decaying States of Plane Strain in a Semi-Infinite Strip and Boundary Conditions for Plate Theory," *J. Elasticity*, Vol. 14, pp. 27–64.

Horgan, C. O., 1982, "Saint-Venant End Effects in Composites," *J. Composite Materials*, Vol. 16, pp. 411–422.

Kapania, R. K., and Raciti, S., 1989, "Recent Advances in Analysis of Laminated Beams and Plates, Part I: Shear Effects and Buckling; Part II: Vibrations and Wave Propagation," *AIAA J.*, Vol. 27, pp. 923–946.

Libai, A., and Simmonds, J. G., 1988, *The Nonlinear Theory of Elastic Shells: One Spatial Dimension*, Academic Press, Boston.

Noor, A. K., and Burton, W. S., 1989, "Assessment of Shear Deformation Theories for Multilayered Composite Plates," *Applied Mechanics Reviews*, Vol. 41, ASME, New York, pp. 1–13.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Equations of Elastic Deformation of Laminated Composite Shallow Shells

A. W. Leissa
Department of Engineering Mechanics,
The Ohio State University,
Columbus, OH 43210
Fellow ASME

M. S. Qatu
Dresser Industries,
Jeffery Division,
Columbus, OH 43201

Shallow shells of laminated composite materials are being increasingly used in structural applications. A complete and consistent theory is needed to deal with elastic deformation problems (i.e., static deflections and stresses, free and forced vibrations). The present work develops equations of motion, which may be solved either exactly or by an approximate method (e.g., Galerkin, finite differences) and energy functionals which may be used with the Ritz or finite element methods to obtain approximate solutions. The equations are developed in terms of arbitrarily-oriented (i.e., nonprincipal) shell coordinates, including twist as well as radii of curvature. The equations account for arbitrary layer thicknesses, fiber orientations, and stacking sequences. Shear deformation and rotary inertia effects are neglected.

Introduction

During the past two decades there has been continually increasing usage of laminated composite materials in structural applications. Often encountered among these applications are plate and shell structural components. Accompanying this increasing usage has been a growth in the literature of composite laminate structural analysis, particularly for plates and, to a lesser extent, for circular cylindrical shells. For example, approximately 300 references dealing with buckling of laminated composite plates were available through 1984 (Leissa, 1985, 1987). Equations are thoroughly developed for the deformation analysis of laminated composite plates (cf., Ambartsumyan, 1964, 1970; Ashton and Whitney, 1970), as well as for circular cylindrical shells.

Shallow shells of laminated construction are also being more frequently used. They can have various types of curvature (e.g., circular cylindrical, spherical, ellipsoidal, hyperbolic paraboloidal). A complete and consistent theory is needed to deal with elastic deformation problems, (i.e., static deflections and stresses, and free and forced vibrations). This requires equations of equilibrium or motion, which may be solved either exactly or by an approximate method (e.g., Galerkin, finite differences), and energy functionals which may be used with the Ritz or finite element methods to obtain approximate solutions. Moreover, the equations of motion and energy functionals should be consistent. That is, they should give the same results for the same problem if solutions are carried out with sufficient numerical accuracy.

Relatively little has been done with composite shallow shells having curvatures other than circular cylindrical. The authors have not found in the published literature a shallow shell theory with sufficient completeness and consistency to deal with arbitrary lamination sequences and arbitrary curvatures. The most useful work to date appears to be the classic monograph by Ambartsumyan (1964), in which one chapter is devoted to laminated composite shallow shells. Another work by Librescu (1975), in which the primary emphasis is aeroelasticity and flutter, also deals extensively with such structural components. However, publications to date appear to be limited to equations of motion expressed in terms of the principal curvatures of shells, and consistent energy functionals also appear to be lacking. Thus, for example, application of the Ritz method to twisted plates representing turbomachinery blades (Leissa, Lee, and Wang, 1984) cannot be undertaken straightforwardly for composite materials.

The objective of the present work is to develop a complete and consistent set of equations which may be used for the stress and deflection analysis of statically or dynamically loaded, laminated composite shallow shells or for their free vibration analyses. The theory will employ the usual assumptions made for laminates; that is, that the layers are each made of parallel fibers embedded in a matrix material, and that each layer may be consequently regarded as being a macroscopically homogeneous, orthotropic, linearly elastic material. The present development is limited to shells sufficiently thin that the effects of shear deformation and rotary inertia may be neglected.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 26, 1989; final revision, Jan. 25, 1990.

Paper No. 91-APM-25.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

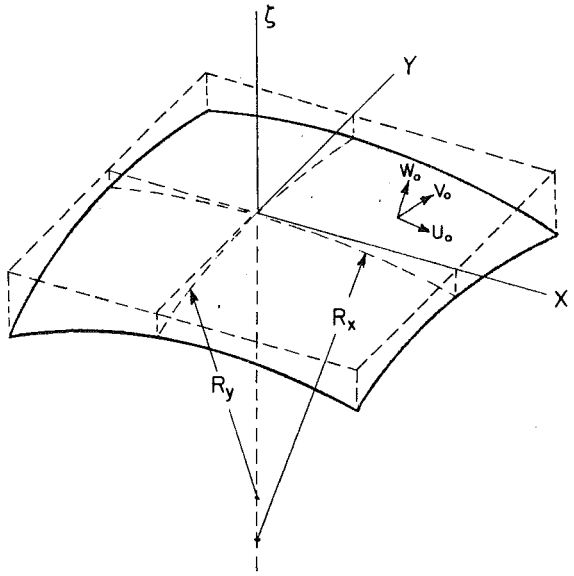


Fig. 1 Middle surface of a shallow shell

Kinematical Relationships

A shallow shell is characterized by its middle surface, which is defined by

$$\zeta = -\frac{1}{2}\left(\frac{x^2}{R_x} + 2\frac{xy}{R_{xy}} + \frac{y^2}{R_y}\right). \tag{1}$$

The constants R_x, R_y, and R_{xy} identify the radii of curvature in the x and y directions and the radius of twist, respectively. A typical quadric middle surface donated by equation (1) is depicted in Fig. 1 for positive values of R_x, R_y, and R_{xy}.

As is customary in shell analysis, an orthogonal set of coordinates α and β are chosen on the middle surface so that the square of the differential of the arc length measured along the surface is given by

$$ds^2 = A^2 d\alpha^2 + B^2\, d\beta^2 \tag{2}$$

where A and B are the Lamé parameters of the surface. Unlike most analyses, the present work will not use principal coordinates (for which $R_{xy} = 0$).

Suppose the shell undergoes deformation so that the displacement of a typical point on the middle surface is given by the set of orthogonal components u_o, v_o, and w_o, where u_o and v_o are tangent to the surface in the directions of α and β, respectively, and w_o is normal to the surface (see Fig. 1).

The middle surface normal strains ϵ^o_α and ϵ^o_β, and the engineering shear strain $\tau^{\ o}_{\alpha\beta}$, are found to be (cf., Gol'denveizer, (1961):

$$\epsilon^o_\alpha = \frac{1}{A}\frac{\partial u_o}{\partial \alpha} + \frac{v_o}{AB}\frac{\partial A}{\partial \beta} + \frac{w_o}{R_\alpha}$$

$$\epsilon^o_\beta = \frac{1}{B}\frac{\partial v_o}{\partial \beta} + \frac{u_o}{AB}\frac{\partial B}{\partial \alpha} + \frac{w_o}{R_\beta}$$

$$\gamma^o_{\alpha\beta} = \frac{B}{A}\frac{\partial}{\partial \alpha}\left(\frac{v_o}{B}\right) + \frac{A}{B}\frac{\partial}{\partial \beta}\left(\frac{u_o}{A}\right) + \frac{2w_o}{R_{\alpha\beta}}. \tag{3}$$

The changes in curvature κ_α and κ_β and in twist τ are:

$$\kappa_\alpha = \frac{1}{A}\frac{\partial}{\partial \alpha}\left(\frac{u_o}{R_\alpha} + \frac{v_o}{R_{\alpha\beta}} - \frac{1}{A}\frac{\partial w_o}{\partial \alpha}\right) + \frac{1}{AB}\frac{\partial A}{\partial \beta}\left(\frac{u_o}{R_{\alpha\beta}} + \frac{v_o}{R_\beta} - \frac{1}{B}\frac{\partial w_o}{\partial \beta}\right) - \frac{1}{2AB\,R_{\alpha\beta}}\left[\frac{\partial}{\partial \beta}(Au_o) - \frac{\partial}{\partial \alpha}(Bv_o)\right]$$

$$\kappa_\beta = \frac{1}{B}\frac{\partial}{\partial \beta}\left(\frac{u_o}{R_{\alpha\beta}} + \frac{v_o}{R_\beta} - \frac{1}{B}\frac{\partial w_o}{\partial \beta}\right) + \frac{1}{AB}\frac{\partial B}{\partial \alpha}\left(\frac{u_o}{R_\alpha} + \frac{v_o}{R_{\alpha\beta}} - \frac{1}{A}\frac{\partial w_o}{\partial \alpha}\right) - \frac{1}{2AB\,R_{\alpha\beta}}\left[\frac{\partial}{\partial \alpha}(Bv_o) - \frac{\partial}{\partial \beta}(Au_o)\right]$$

$$\frac{\tau}{2} = \frac{1}{A}\frac{\partial}{\partial \alpha}\left(\frac{u_o}{R_{\alpha\beta}} + \frac{v_o}{R_\beta} - \frac{1}{B}\frac{\partial w_o}{\partial \beta}\right) - \frac{1}{AB}\frac{\partial A}{\partial \beta}\left(\frac{u_o}{R_\alpha} + \frac{v_o}{R_{\alpha\beta}} - \frac{1}{A}\frac{\partial w_o}{\partial \alpha}\right) + \frac{1}{R_\alpha}\left(\frac{1}{B}\frac{\partial u_o}{\partial \beta} - \frac{v_o}{AB}\frac{\partial B}{\partial \alpha} + \frac{w_o}{R_{\alpha\beta}}\right) + \frac{1}{R_{\alpha\beta}}\left(\frac{1}{B}\frac{\partial v_o}{\partial \beta} + \frac{u_o}{AB}\frac{\partial B}{\partial \alpha} + \frac{w_o}{R_\beta}\right). \tag{4}$$

Equations (3) and (4) may be transformed from the shell coordinates α, β to the planform coordinates x, y, which are more convenient for shallow shells (see Fig. 1). Further, the radii of curvature R_α, R_β, and $R_{\alpha\beta}$ are constant in the present problem, and correspond to R_x, R_y, and R_{xy} of equation (1), and $A = B = 1$. Finally, it is customarily assumed that the curvature changes caused by the tangential displacement components (u_o, v_o) are small in a shallow shell, in comparison with changes caused by the normal component (w_o) (cf., Ambartsumyan, (1964), and that the terms $w_o/R_\alpha R_{\alpha\beta}$ and $w_o/R_\beta R_{\alpha\beta}$ in the last of the equations (4) are relatively small. With these considerations, equations (3) and (4) become

$$\epsilon^o_x = \frac{\partial u_o}{\partial x} + \frac{w_o}{R_x}$$

$$\epsilon^o_y = \frac{\partial v_o}{\partial y} + \frac{w_o}{R_y}$$

$$\gamma^o_{xy} = \frac{\partial v_o}{\partial x} + \frac{\partial u_o}{\partial y} + \frac{2w_o}{R_{xy}} \tag{5}$$

$$\kappa_x = -\frac{\partial^2 w_o}{\partial x^2}, \quad \kappa_y = -\frac{\partial^2 w_o}{\partial y^2}, \quad \tau = -2\frac{\partial^2 w_o}{\partial x \partial y}. \tag{6}$$

To complete the needed kinematical relationships, it is also necessary to determine the strain components at an *arbitrary* point located within the shell (i.e., not necessarily at the middle surface). Utilizing the well-known Kirchhoff hypothesis (normals to the middle surface remain straight and normal, and unstretched in length deformation (cf., Love, 1927), the displacements of an arbitrary point are:

$$u = u_o - z\frac{\partial w_o}{\partial \alpha}$$

$$v = v_o - z\frac{\partial w_o}{\partial \beta}$$

$$w = w_o \tag{7}$$

where z is the coordinate measured from the middle surface in the direction of the outer normal. Expressing equations (7) in terms of the planform coordinates,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$u = u_o - z\frac{\partial w_o}{\partial x}$$

$$v = v_o - z\frac{\partial w_o}{\partial y}$$

$$w = w_o. \tag{8}$$

The strain components at a typical point are

$$\epsilon_x = \frac{\partial u}{\partial x} + \frac{w_o}{R_x + z}$$

$$\epsilon_y = \frac{\partial v}{\partial y} + \frac{w_o}{R_y + z}$$

$$\gamma_{xy} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} + \frac{2w_o}{R_{xy} + z}. \tag{9}$$

Substituting equations (8) into (9), assuming that z is negligible in comparison with R_x, R_y, and R_{xy}, there results:

$$\epsilon_x = \epsilon_x^o + z\,\kappa_x$$

$$\epsilon_y = \epsilon_y^o + z\,\kappa_y \tag{10}$$

$$\gamma_{xy} = \gamma_{xy}^o + z\tau.$$

Force and Moment Resultants

Consider now a laminated shallow shell composed of composite material laminae (typically, very thin layers). The materials of each lamina consists of parallel, continuous fibers of one material (e.g., glass, boron, graphite) embedded in a matrix material (e.g., epoxy resin), the primary purpose of which is to transfer shear stress between the fibers as needed. Each layer may be regarded on a macroscopic scale as being homogeneous and rectangularly orthotropic. However, the fibers of a typical layer may not be parallel to the coordinates in which the shell equations are expressed. Then the stress-strain equations for an element of material in the kth lamina may be written as (cf. Jones, 1975; Vinson and Sierakowski, 1986):

$$\begin{bmatrix} \sigma_\alpha \\ \sigma_\beta \\ \sigma_z \\ \tau_{\beta z} \\ \tau_{\alpha z} \\ \tau_{\alpha\beta} \end{bmatrix}_k = \begin{bmatrix} \overline{Q}_{11} & \overline{Q}_{12} & \overline{Q}_{13} & 0 & 0 & \overline{Q}_{16} \\ \overline{Q}_{12} & \overline{Q}_{22} & \overline{Q}_{23} & 0 & 0 & \overline{Q}_{26} \\ \overline{Q}_{13} & \overline{Q}_{23} & \overline{Q}_{33} & 0 & 0 & \overline{Q}_{36} \\ 0 & 0 & 0 & \overline{Q}_{44} & \overline{Q}_{45} & 0 \\ 0 & 0 & 0 & \overline{Q}_{45} & \overline{Q}_{55} & 0 \\ \overline{Q}_{16} & \overline{Q}_{26} & \overline{Q}_{36} & 0 & 0 & \overline{Q}_{66} \end{bmatrix} \begin{bmatrix} \epsilon_\alpha \\ \epsilon_\beta \\ \epsilon_z \\ \gamma_{\beta z} \\ \gamma_{\alpha z} \\ \gamma_{\alpha\beta} \end{bmatrix}_k \tag{11}$$

where σ_α, σ_β, and σ_z are normal components; $\tau_{\alpha z}$, $\tau_{\beta z}$, and $\tau_{\alpha\beta}$ are shear stress components (see Fig. 2); and $\epsilon_{\alpha z}$, $\epsilon_{\beta z}$, and $\epsilon_{\alpha\beta}$ are the *tensorial* shear strains ($\epsilon_{\alpha\beta} = \gamma_{\alpha\beta}/2$, etc.). The constants $\overline{Q}_{ij}$ are the elastic stiffness coefficients for the material, which are determined from a micromechanical analysis of the material, considering the elastic properties of the fiber and matrix materials and the percent fiber by volume.

Applying the Kirchoff hypothesis and the assumption that σ_z is negligible, as is customarily made in thin shell theory (cf., Love, 1927), equations (11) reduce to

$$\begin{bmatrix} \sigma_\alpha \\ \sigma_\beta \\ \tau_{\alpha\beta} \end{bmatrix}_k = \begin{bmatrix} \overline{Q}_{11} & \overline{Q}_{12} & \overline{Q}_{16} \\ \overline{Q}_{12} & \overline{Q}_{22} & \overline{Q}_{26} \\ \overline{Q}_{16} & \overline{Q}_{26} & \overline{Q}_{66} \end{bmatrix}_k \begin{bmatrix} \epsilon_\alpha \\ \epsilon_\beta \\ \gamma_{\alpha\beta} \end{bmatrix}_k. \tag{12}$$

Although equation (12) appears to be for a generally anisotropic material, the $\overline{Q}_{16}$ and $\overline{Q}_{26}$ coefficients result from a coordinate transformation, and arise only if the shell coordinates (α, β) are not parallel and perpendicular to the fibers. Otherwise, $\overline{Q}_{16}$ and $\overline{Q}_{26}$ are zero.

The force resultants are the integrals of the stresses over the shell thickness (h) (Gol'denveizer, 1961):

$$\begin{bmatrix} N_\alpha \\ N_{\alpha\beta} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_\alpha \\ \tau_{\alpha\beta} \end{bmatrix}_k \left(1 + \frac{z}{R_\beta}\right) dz$$

$$\begin{bmatrix} N_\beta \\ N_{\beta\alpha} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_\beta \\ \tau_{\beta\alpha} \end{bmatrix}_k \left(1 + \frac{z}{R_\alpha}\right) dz. \tag{13}$$

They have dimensions of force per unit length. The moment resultants are the bending moments per unit length:

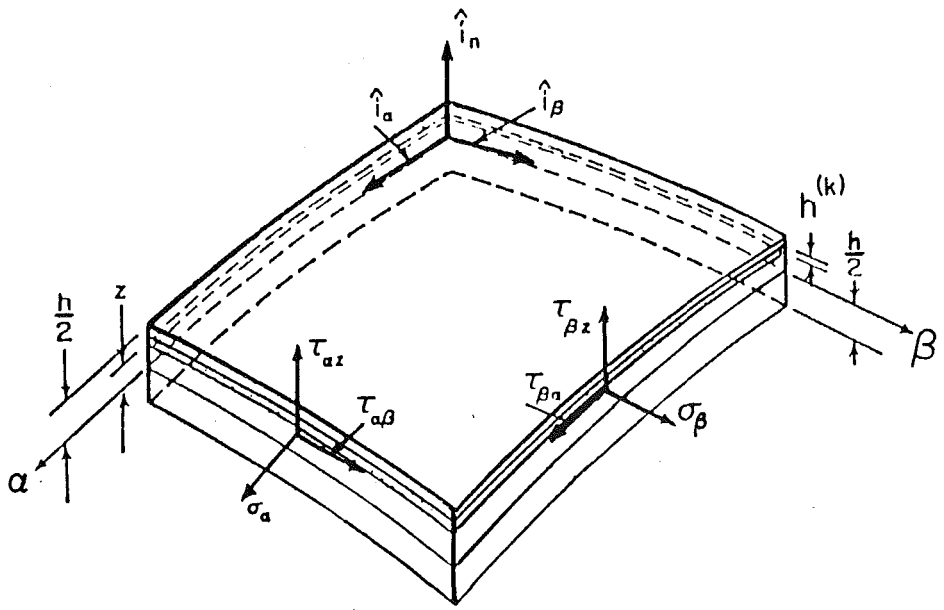


Fig. 2 Stress components with positive directions, in shell coordinates

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\begin{bmatrix} M_\beta \\ M_{\alpha\beta} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_\alpha \\ \tau_{\alpha\beta} \end{bmatrix}_k z \left(1 + \frac{z}{R_\beta}\right) dz$$

$$\begin{bmatrix} M_\beta \\ M_{\beta\alpha} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_\beta \\ \tau_{\beta\alpha} \end{bmatrix}_k z \left(1 + \frac{z}{R_\alpha}\right) dz. \tag{14}$$

Although $\tau_{\alpha\beta} = \tau_{\beta\alpha}$, equations (13) and (14) indicate that $N_{\alpha\beta} \neq N_{\beta\alpha}$, and $M_{\alpha\beta} \neq M_{\beta\alpha}$, unless $R_\alpha = R_\beta$. This would occur, for example, in the case of a spherical shell ($R_\alpha = R_\beta = R$, $R_{\alpha\beta} = 0$) or a twisted plate ($R_\alpha = R_\beta = 0$, $R_{\alpha\beta} \neq 0$).

For a thin shell, the terms z/R_α and z/R_β are much less than unity and may therefore be neglected in equations (13) and (14). Doing so, then $N_{\alpha\beta} = N_{\beta\alpha}$ and $M_{\alpha\beta} = M_{\beta\alpha}$. Introducing the planform coordinates x and y, equations (13) and (14) may be rewritten as:

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} dz \tag{15}$$

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \int_{-\frac{h}{2}}^{\frac{h}{2}} \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} z\, dz. \tag{16}$$

Substituting equation (12) into (15) and (16), utilizing equations (10), and carrying out the integration over the thickness piecewise, from layer to layer,

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \\ \text{---} \\ M_x \\ M_y \\ M_{xy} \end{bmatrix} = \left[\begin{array}{ccc:ccc} A_{11} & A_{12} & A_{16} & B_{11} & B_{12} & B_{16} \\ A_{12} & A_{22} & A_{26} & B_{12} & B_{22} & B_{26} \\ A_{16} & A_{26} & A_{66} & B_{16} & B_{26} & B_{66} \\ \hdashline B_{11} & B_{12} & B_{16} & D_{11} & D_{12} & D_{16} \\ B_{12} & B_{22} & B_{26} & D_{12} & D_{22} & D_{26} \\ B_{16} & B_{26} & B_{66} & D_{16} & D_{26} & D_{66} \end{array}\right] \begin{bmatrix} \epsilon_x^o \\ \epsilon_y^o \\ \gamma_{xy}^o \\ \text{---} \\ \kappa_x \\ \kappa_y \\ \tau \end{bmatrix} \tag{17}$$

where the A_{ij}, B_{ij}, and D_{ij} are the stiffness coefficients arising from the piecewise integration:

$$A_{ij} = \sum_{k=1}^{N} \bar{Q}_{ij}^{(k)} (z_k - z_{k-1})$$

$$B_{ij} = \frac{1}{2} \sum_{k=1}^{N} \bar{Q}_{ij}^{(k)} (z_k^2 - z_{k-1}^2)$$

$$D_{ij} = \frac{1}{3} \sum_{k=1}^{N} \bar{Q}_{ij}^{(k)} (z_k^3 - z_{k-1}^3) \tag{18}$$

where z_k is the distance from the midsurface to the surface of the kth layer having the largest z-coordinate (cf., Jones, 1975). For shells which are laminated symmetrically with respect to their midsurfaces, the B_{ij} are zero. Equations (18) include terms involving the individual layer stiffness coefficient $\bar{Q}_{ij}^{(k)}$ and thickness h_k, summed over the entire N layers of the laminated shell.

Equations (17) and (18) are the same as those for laminated plates (cf., Ambartsumyan, 1970; Whitney, 1987; Jones, 1975). This is a consequence of neglecting terms of the type z/R_α in comparison with unity in equations (9) and (14). this is also made in many deep shell theories, but not all of them (e.g., Flügge, 1962; Lur'ye, 1940).

Equations of Motion

Equations of motion may be obtained by taking a differential element of shell having thickness h and midsurface lengths $Ad\alpha$ by $Bd\beta$, and requiring the sum of the external and internal forces and the sum of the external moments to each be zero.

Figure 3 shows such an element (although the thickness is not described) with all the internal forces (per unit length) acting upon it. These include the force resultants N_α, N_β, $N_{\alpha\beta}$, $N_{\beta\alpha}$ tangent to the midsurface (also called "membrane force resultants") described earlier, and the transverse shearing force resultants Q_α and Q_β. Summing forces on the element yields (Gol'denveizer, 1961):

$$\frac{\partial}{\partial\alpha}(BN_\alpha) + \frac{\partial}{\partial\beta}(AN_{\beta\alpha}) + \frac{\partial A}{\partial\beta} N_{\alpha\beta} - \frac{\partial B}{\partial\alpha} N_\beta + \frac{AB}{R_\alpha} Q_\alpha + \frac{AB}{R_{\alpha\beta}} Q_\beta + ABp_\alpha = \bar{\rho} \frac{\partial^2 u}{\partial t^2}$$

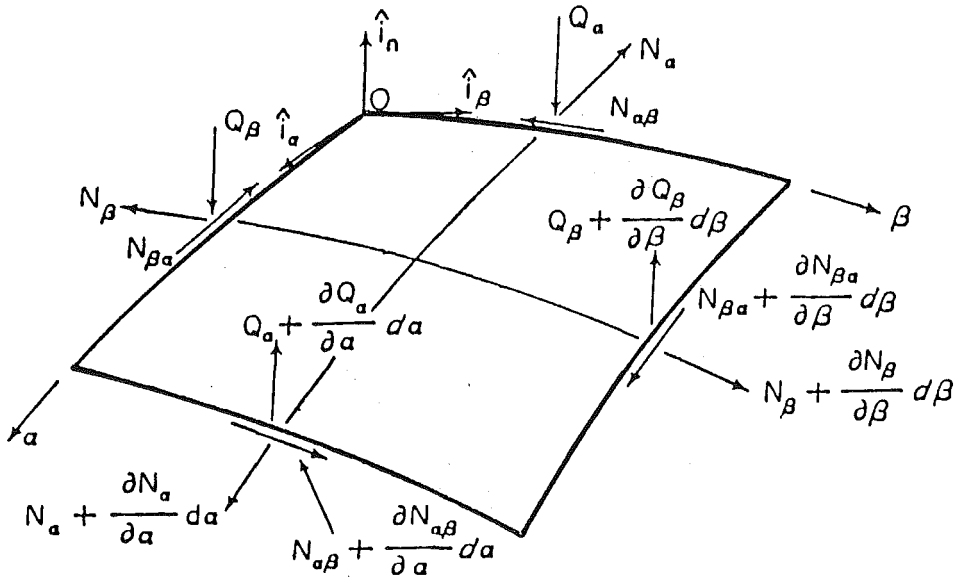


Fig. 3 Force resultants in shell coordinates

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

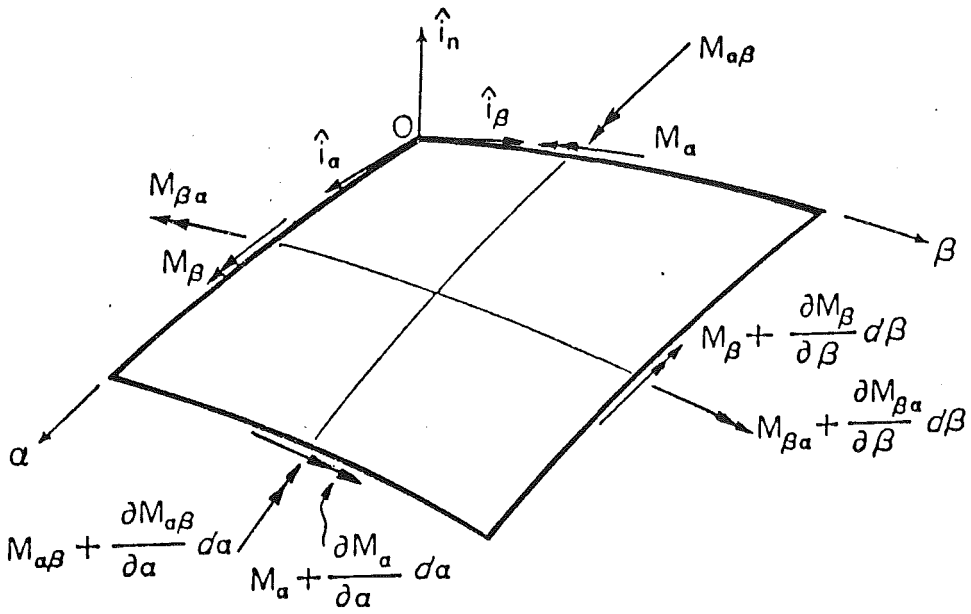


Fig. 4 Moment resultants in shell coordinates

$$\frac{\partial}{\partial\alpha}(BN_\alpha)+\frac{\partial}{\partial\beta}(AN_{\beta\alpha})+\frac{\partial A}{\partial\beta}N_{\alpha\beta}-\frac{\partial B}{\partial\alpha}N_\beta+\frac{AB}{R_\alpha}Q_\alpha + \frac{AB}{R_{\alpha\beta}}Q_\alpha + ABp_\beta=\bar{\rho}\frac{\partial^2 v}{\partial t^2}$$

$$-\frac{AB}{R_{\alpha\beta}}(N_{\alpha\beta}+N_{\beta\alpha})-\frac{AB}{R_\alpha}N_\alpha-\frac{AB}{R_\beta}+\frac{\partial}{\partial\alpha}(BQ_\alpha) + \frac{\partial}{\partial\beta}(AQ_\beta)+ABp_n=\bar{\rho}\frac{\partial^2 w}{\partial t^2} \tag{19}$$

where p_α, p_β, and p_n represent external forces (per unit area) applied either to the surface or as body forces, and where $\bar{\rho}$ is the average mass density of the shell per unit midsurface area, given by

$$\bar{\rho}=\sum_{k=1}^{N}\rho^{(k)}(z_k-z_{k-1}). \tag{20}$$

Figure 4 depicts the element with the moment resultants acting upon it. Summing moments on the element, including both the moment resultants and the force resultants gives:

$$\frac{\partial}{\partial\alpha}(BM_\alpha)+\frac{\partial}{\partial\beta}(AM_{\beta\alpha})+\frac{\partial A}{\partial\beta}M_{\alpha\beta}-\frac{\partial B}{\partial\alpha}M_\beta - AB\,Q_\alpha+AB\,m_\beta=0$$

$$\frac{\partial}{\partial\beta}(AM_\beta)+\frac{\partial}{\partial\alpha}(BM_{\alpha\beta})+\frac{\partial B}{\partial\alpha}M_{\beta\alpha}-\frac{\partial A}{\partial\beta}M_\alpha - AB\,Q_\beta+AB\,m_\alpha=0$$

$$N_{\alpha\beta}-N_{\beta\alpha}+\frac{M_{\alpha\beta}}{R_\alpha}-\frac{M_{\beta\alpha}}{R_\beta}+\frac{1}{R_{\alpha\beta}}(M_\beta-M_\alpha)=0 \tag{21}$$

where m_α and m_β represent possible body couples (moments per unit length) which may be applied to the shell. Rotary inertia terms are omitted from the right-hand sides of equations (21) because, as it is well known, they are less important than shear deformation effects, which are being neglected in the present analysis.

The equations of motion for the shallow shell are obtained from equations (19) and (21) by changing from the α, β shell coordinates to the x, y planform coordinates, setting R_α, R_β, $R_{\alpha\beta}$ equal to R_x, R_y, and R_{xy}, and $A=B=1$. Furthermore, for a shallow shell, $N_{xy}=N_{yx}$ and $M_{xy}=M_{yx}$. The equation of motion therefore becomes:

$$\frac{\partial N_x}{\partial x}+\frac{\partial N_{xy}}{\partial y}+\frac{Q_x}{R_x}+\frac{Q_y}{R_{xy}}=\bar{\rho}\frac{\partial^2 u}{\partial t^2}-p_x \tag{22a}$$

$$\frac{\partial N_{xy}}{\partial x}+\frac{\partial N_y}{\partial y}+\frac{Q_x}{R_{xy}}+\frac{Q_y}{R_y}=\bar{\rho}\frac{\partial^2 v}{\partial t^2}-p_y \tag{22b}$$

$$-\frac{N_x}{R_x}-\frac{N_y}{R_y}-\frac{2N_{xy}}{R_{xy}}+\frac{\partial Q_x}{\partial x}+\frac{\partial Q_y}{\partial y}=\bar{\rho}\frac{\partial^2 w}{\partial t^2}-p_n \tag{22c}$$

$$\frac{\partial M_x}{\partial x}+\frac{\partial M_{xy}}{\partial y}-Q_x=0 \tag{22d}$$

$$\frac{\partial M_{xy}}{\partial x}+\frac{\partial M_y}{\partial y}-Q_y=0 \tag{22e}$$

where p_x and p_y are external force components tangent to the shell midsurface. The last of equations (21) is typically ignored in shell analysis. Although it may not be identically satisfied, the nonzero residual that will be found on its right-hand side after the solution is substituted into it, which may be interpreted as a body moment about the n-axis, will typically be small.

For a shallow shell the terms involving Q_z and Q_y in equations (22a) and (22b) are typically small in comparison with the remaining terms, and are neglected. Solving equations (22d) and (22e) for Q_x and Q_y, and substituting these into equation (22c), the equations of motion become:

$$\frac{\partial N_x}{\partial x}+\frac{\partial N_{xy}}{\partial y}=\bar{\rho}\frac{\partial^2 u}{\partial t^2}-p_x$$

$$\frac{\partial N_{xy}}{\partial x}+\frac{\partial N_y}{\partial y}=\bar{\rho}\frac{\partial^2 v}{\partial t^2}-p_y$$

$$-\frac{N_x}{R_x}-\frac{N_y}{R_y}-\frac{2\partial N_{xy}}{R_{xy}}+\frac{\partial^2 M_x}{\partial x^2}+2\frac{\partial^2 M_{xy}}{\partial x\partial y}+\frac{\partial^2 M_y}{\partial y^2}=\bar{\rho}\frac{\partial^2 w}{\partial t^2}-p_n. \tag{23}$$

Multiplying the last of equations (23) through by -1 and substituting equations (5), (6), and (17), the equations of mo-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

tion may be expressed in terms of the midsurface displacement components in matrix form as:

$$\begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} u_o \\ v_o \\ w_o \end{bmatrix} + \begin{bmatrix} -\bar{\rho} & 0 & 0 \\ 0 & -\bar{\rho} & 0 \\ 0 & 0 & \bar{\rho} \end{bmatrix} \frac{\partial^2}{\partial t^2} \begin{bmatrix} u_o \\ v_o \\ w_o \end{bmatrix} = \begin{bmatrix} -p_x \\ -p_y \\ p_n \end{bmatrix} \tag{24}$$

where

$$L_{11} = A_{11}\frac{\partial^2}{\partial x^2} + 2A_{16}\frac{\partial^2}{\partial x \partial y} + A_{66}\frac{\partial^2}{\partial y^2}$$

$$L_{12} = L_{21} = A_{16}\frac{\partial^2}{\partial x^2} + (A_{12}+A_{66})\frac{\partial^2}{\partial x \partial y} + A_{26}\frac{\partial^2}{\partial y^2}$$

$$L_{22} = A_{66}\frac{\partial^2}{\partial x^2} + 2A_{26}\frac{\partial^2}{\partial x \partial y} + A_{22}\frac{\partial^2}{\partial y^2}$$

$$L_{13} = L_{31} = -\left\{B_{11}\frac{\partial^3}{\partial x^3} + 3B_{16}\frac{\partial^3}{\partial x^2 \partial y} + (B_{12}+2B_{66})\frac{\partial^3}{\partial x \partial y^2} + B_{26}\frac{\partial^3}{\partial y^3}\right\} + \left\{\frac{A_{11}}{R_x} + \frac{A_{12}}{R_y} + 2\frac{A_{16}}{R_{xy}}\right\}\frac{\partial}{\partial x} + \left\{\frac{A_{16}}{R_x} + \frac{A_{26}}{R_y} + 2\frac{A_{66}}{R_{xy}}\right\}\frac{\partial}{\partial y}$$

$$L_{23} = L_{32} = -\left\{B_{16}\frac{\partial^3}{\partial x^3} + 3B_{26}\frac{\partial^3}{\partial x \partial y^2} + (B_{12}+2B_{66})\frac{\partial^3}{\partial x^2 \partial y} + B_{22}\frac{\partial^3}{\partial y^3}\right\} + \left\{\frac{A_{12}}{R_x} + \frac{A_{22}}{R_y} + 2\frac{A_{26}}{R_{xy}}\right\}\frac{\partial}{\partial y} + \left\{\frac{A_{16}}{R_x} + \frac{A_{26}}{R_y} + 2\frac{A_{66}}{R_{xy}}\right\}\frac{\partial}{\partial x}$$

$$L_{33} = \left\{D_{11}\frac{\partial^4}{\partial x^4} + 4D_{16}\frac{\partial^4}{\partial x^3 \partial y} + 2(D_{12}+2D_{66})\frac{\partial^4}{\partial x^2 \partial y^2} + 4D_{26}\frac{\partial^4}{\partial x \partial y^3} + D_{22}\frac{\partial^4}{\partial y^4}\right\} - 2\left\{\left(\frac{B_{11}}{R_x} + \frac{B_{12}}{R_y} + 2\frac{B_{16}}{R_{xy}}\right)\frac{\partial^2}{\partial x^2} + 2\left(\frac{B_{16}}{R_x} + \frac{B_{26}}{R_y} + 2\frac{B_{66}}{R_{xy}}\right)\frac{\partial^2}{\partial x \partial y} + \left(\frac{B_{12}}{R_x} + \frac{B_{22}}{R_y} + 2\frac{B_{26}}{R_{xy}}\right)\frac{\partial^2}{\partial y^2}\right\} + \left\{\frac{A_{11}}{R_x^2} + 2\frac{A_{12}}{R_x R_y} + \frac{A_{22}}{R_y^2} + 4\left(\frac{A_{16}}{R_{xy}R_x} + \frac{A_{26}}{R_{xy}R_y} + \frac{A_{66}}{R_{xy}^2}\right)\right\}. \tag{25}$$

For symmetrically laminated shells, terms containing B_{ij} all vanish.

Boundary Conditions

Boundary conditions for a shallow shell may be derived by physical arguments, or by using energy functionals and the principle of virtual work. On each boundary one must specify four conditions. If n and t represent the directions normal and tangent to the boundary in the plane tangent to the midsurface at a point, the four "simple" boundary conditions which may be encountered for a shell are (cf., Gol'denveizer, 1961)

$$u_n = 0 \quad \text{or} \quad N_n = 0 \tag{26a}$$

$$u_t = 0 \quad \text{or} \quad N_{nt} = 0 \tag{26b}$$

$$w = 0 \quad \text{or} \quad Q_n + \frac{\partial M_{nt}}{\partial t} = 0 \tag{26c}$$

$$\frac{\partial w}{\partial n} = 0 \quad \text{or} \quad M_n = 0 \tag{26d}$$

where u_n and u_t are displacements in the normal and tangential directions, respectively. In addition, at any corner on the boundary,

$$w = 0 \quad \text{or} \quad M_{nt} = 0. \tag{27}$$

It should be noted that the second of the two possible boundary conditions in equation (26*b*), which will occur at a free edge, does not contain a term M_{nt}/R_t as it would in the case of a deep shell. This term is negligible to the same order of approximation made elsewhere in the shallow shell theory; viz., neglecting curvature changes due to tangential displacement components. A variational derivation also shows that $N_{nt} = 0$ is a proper boundary condition (Qatu, 1989).

Boundaries may also be elastically constrained, with the constraints being represented as translational and rotational springs distributed along the boundaries. In such cases the boundary conditions are generalized to

$$k_u u + N_n = 0$$

$$k_v v + N_{nt} = 0$$

$$k_w w + Q_n + \frac{\partial M_{nt}}{\partial t} = 0$$

$$k_\psi \frac{\partial w}{\partial n} - M_n = 0 \tag{28}$$

along the boundaries, where k_u, k_v, k_w are distributed translational spring stiffnesses and k_ψ is the distributed rotational spring stiffness. At a corner the generalized condition becomes

$$K_w w - 2M_{nt} = 0 \tag{29}$$

where K_w is the stiffness of a translational spring concentrated at the corner.

Energy Functionals

The strain energy stored in a shell during elastic deformation is

$$U = \frac{1}{2}\int_v (\sigma_\alpha \epsilon_\alpha + \sigma_\beta \epsilon_\beta + \tau_{\alpha\beta}\gamma_{\alpha\beta}) dV \tag{30}$$

where V is the volume. For the kth lamina, using equation (12), this becomes

$$U_k = \frac{1}{2}\int_{A_k}\int_{z_{k-1}}^{z_k} (\bar{Q}_{11}\epsilon_\alpha^2 + \bar{Q}_{22}\epsilon_\beta^2 + \bar{Q}_{66}\gamma_{\alpha\beta}^2 + 2\bar{Q}_{12}\epsilon_\alpha\epsilon_\beta + 2\bar{Q}_{16}\epsilon_\alpha\gamma_{\alpha\beta} + 2\bar{Q}_{26}\epsilon_\beta\gamma_{\alpha\beta}) dz\, dA_k. \tag{31}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Thus, for the complete shell,

$$U=\sum_{k=1}^{N} U_k \tag{32}$$

where again N is the total number of laminae.

Transforming to the planform coordinates x and y, substituting equations (10), and carrying out the integration through the thickness, the strain energy of the thin shell becomes

$$\begin{aligned}U=\frac{1}{2}\int_A\int\Big\{&\Big[A_{11}(\epsilon_x^o)^2+2B_{11}\epsilon_x^o\kappa_x+D_{11}\kappa_x^2\Big]\\&+\Big[A_{22}(\epsilon_y^o)^2+2B_{22}\epsilon_y^o\kappa_y+D_{22}\kappa_y^2\Big]\\&+\Big[A_{66}(\gamma_{xy}^o)^2+2B_{66}\gamma_{xy}^o\tau+D_{66}\tau^2\Big]\\&+2\Big[A_{12}\epsilon_x^o\epsilon_y^o+B_{12}(\epsilon_x^o\kappa_y+\epsilon_y^o\kappa_x)+D_{12}\kappa_x\kappa_y\Big]\\&+2\Big[A_{16}\epsilon_x^o\gamma_{xy}^o+B_{16}(\epsilon_x^o\tau+\gamma_{xy}^o\kappa_x)+D_{16}\kappa_x\tau\Big]\\&+2\Big[A_{26}\epsilon_y^o\gamma_{xy}^o+B_{26}(\epsilon_y^o\tau+\gamma_{xy}^o\kappa_y)+D_{26}\kappa_y\tau\Big]\Big\}dA.\end{aligned} \tag{33}$$

Substituting the strain-displacement and curvature-displacement equations (5) and (6) into (33) yields the strain energy in terms of the displacements. This may be expressed as the sum of three parts:

$$U=U_s+U_b+U_{bs} \tag{34}$$

where U_s is the part due to stretching alone,

$$\begin{aligned}U=\frac{1}{2}\int_A\int\Big\{&A_{11}\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)^2+A_{22}\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)^2,\\&+A_{66}\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+\frac{2w_o}{R_{xy}}\right)^2\\&+2A_{12}\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)\\&+2A_{16}\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+\frac{2w_o}{R_{xy}}\right)\\&+2A_{26}\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+\frac{2w_o}{R_{xy}}\right)\Big\}dA,\end{aligned} \tag{35}$$

U_b is the part due to bending alone,

$$\begin{aligned}U_b=\frac{1}{2}\int_A\int\Big\{&D_{11}\left(\frac{\partial^2 w_o}{\partial x^2}\right)^2+D_{22}\left(\frac{\partial^2 w_o}{\partial y^2}\right)+4D_{66}\left(\frac{\partial^2 w_o}{\partial x\partial y}\right)^2\\&+2D_{12}\frac{\partial^2 w_o}{\partial x^2}\frac{\partial^2 w_o}{\partial y^2}+4D_{16}\frac{\partial^2 w_o}{\partial x^2}\frac{\partial^2 w_o}{\partial x\partial y}+4D_{26}\frac{\partial^2 w_o}{\partial y^2}\frac{\partial^2 w_o}{\partial x\partial y}\Big\}dA,\end{aligned} \tag{36}$$

and U_{bs} is the part due to bending-stretching coupling,

$$\begin{aligned}U_{bs}=-\int_A\int\Big\{&B_{11}\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)\frac{\partial^2 w_o}{\partial x^2}+B_{22}\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)\frac{\partial^2 w_o}{\partial y^2}\\&+2B_{66}\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+2\frac{w_o}{R_{xy}}\right)\frac{\partial^2 w_o}{\partial x\partial y}\\&+B_{12}\left[\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)\frac{\partial^2 w_o}{\partial y^2}+\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)\frac{\partial^2 w_o}{\partial x^2}\right]\\&+B_{16}\left[2\left(\frac{\partial u_o}{\partial x}+\frac{w_o}{R_x}\right)\frac{\partial^2 w_o}{\partial x\partial y}+\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+2\frac{w_o}{R_{xy}}\right)\frac{\partial^2 w_o}{\partial x^2}\right]\\&+B_{26}\left[2\left(\frac{\partial v_o}{\partial y}+\frac{w_o}{R_y}\right)\frac{\partial^2 w_o}{\partial x\partial y}+\left(\frac{\partial v_o}{\partial x}+\frac{\partial u_o}{\partial y}+2\frac{w_o}{R_{xy}}\right)\frac{\partial^2 w_o}{\partial y^2}\right]\Big\}dA.\end{aligned} \tag{37}$$

Comparing the strain energy functional of equations (34)–(37) with that derived by Rao (1978), one finds many terms in the latter which are negligible for shallow shells, although that work was intended for shallow shells. Equations (34)–(37) were alos developed by Baharlou (1985). For symmetrically laminated shells. $B_{ij}=0$ and, hence, $U_{bs}=0$.

Using the distributed external force components p_x and p_y in the tangential directions, and p_n in the normal direction, the work done by the external forces as the shell displaces is

$$W=\int_A\int(p_x u_o+p_y v_o+p_n w_o)\,dA. \tag{38}$$

Therefore, the total potential energy of the shell in its deformed shape is

$$\Pi=U-W. \tag{39}$$

The kinetic energy for each lamina is

$$T_k=\frac{1}{2}\rho^{(k)}\int_{A_k}\int_{z_{k-1}}^{z_k}\left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial v}{\partial t}\right)^2+\left(\frac{\partial w}{\partial t}\right)^2\right]dz\,dA_k \tag{40}$$

where $\rho^{(k)}$ is the density of the lamina per unit volume, and t is time. The kinetic energy of the entire shell is then found to be (neglecting rotary inertia terms):

$$T=\sum_{k=1}^{N}T_k=\frac{\bar{\rho}}{2}\int_A\int\left[\left(\frac{\partial u_o}{\partial t}\right)^2+\left(\frac{\partial v_o}{\partial t}\right)^2+\left(\frac{\partial w_o}{\partial t}\right)^2\right]dA \tag{41}$$

where $\bar{\rho}$ is the average mass density of the shell per unit area of the midsurface, as given by equation (20).

Concluding Remarks

A complete and consistent set of equations has been developed which characterizes the static and dynamic deformation of laminated composite shallow shells. The equations account for composite materials arranged in arbitrary layer thicknesses, fiber orientations, and stacking sequences. The shells may have general radii of curvature, although taken to be constant, which permits the analysis of, for example, circular cylindrical, spherical, ellipsoidal, and hyperbolic paraboloidal shallow shells. Initial twist (characterized by R_{xy}) has also been included, which permits one to analyze twisted plates of composite material.

The eighth-order set of differential equations of motion are useful for obtaining exact solutions, where possible, or for approximate analyses employing, for example, Galerkin or finite difference methods. By introducing the stiffness and kinematical relationships into the equations of motion, they have been expressed in terms of displacements. If the assumed displacement functions have sufficient continuity, then the equations of compatibility will be identically satisfied. Many shell analysts prefer to introduce an Airy stress function into the problem, and obtain an eighth-order set of differential equations which includes a compatibility condition. However, the writers see no advantage in this formulation of the problem. Moreover, it is applicable only to static problems. Therefore it has not been included.

The energy functionals derived here may be used with the Ritz method to solve free vibration problems for laminated composite shallow shells. This method has been shown to be very successful for homogeneous, isotropic shallow shells (Leissa, Lee, and Wang, 1983, 1984; Leissa and Narita, 1984) and for laminated composite ones (Qatu, 1989). Moreover,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

the method works well for static or forced vibration problems. For such problems it has been shown that not only displacements, but membrane and bending stresses, may be determined with sufficient accuracy by the Ritz method. The energy functionals have been shown to be consistent with the differential equations and boundary conditions presented (Qatu, 1989). They may also be used in the development of consistent finite element methods.

It has been shown that the stiffness relationship between the force and moment resultants and the midsurface strains and bending curvature changes (equation (17)) for thin shells can be taken in the same form as for plates. However, for thick, laminated composite shells, it would have to be reconsidered, not only to include shear flexibility effects, but to retain terms of the type z/R_x when integrating the stresses over the thickness to obtain force and moment resultants.

References

Ambartsumyan, S. A., 1964, *Theory of Anisotropic Shells*, (English translation), NASA TT-F118, p. 396.

Ambartsumyan, S. A., 1970, *Theory of Anisotropic Plates*, (English translation), Technomic Publishing, Inc., Stamford, Conn.

Baharlou, B., 1985, "Vibration and Buckling of Laminated Composite Plates with Arbitrary Boundary Conditions," Ph.D. dissertation, Ohio State University.

Flügge, W., 1962, *Stresses in Shells*, Springer-Verlag, Berlin.

Gol'denveizer, A. L., 1961, *Theory of Elastic Thin Shells* (English translation), Pergamon Press, New York.

Jones, R. M., 1975, *Mechanics of Composite Materials*, Washington.

Leissa, A. W., 1985, "Buckling of Laminated Composite Plates and Shell Panels," Flight Dynamics Laboratory Report No. AFWAL-TR-85-3069, Wright-Patterson Air Force Base, p. 439.

Leissa, A. W., 1987, "A Review of Laminated Composite Plate Buckling," ASME *Applied Mechanics Reviews*, Vol. 40, pp. 575–591.

Leissa, A. W., Lee, J. K., and Wang, A. J., 1983, "Vibrations of Doubly-Curved Shallow Shells," *International Journal of Solids and Structures*, Vol. 19, No. 5, pp. 411–424.

Leissa, A. W., Lee, J. K., and Wang, A. J., 1984, "Vibrations of Twisted Rotating Blades," *ASME Journal of Vibration, Acoustics, Stress and Reliability in Design*, Vol. 106, pp. 251–257.

Leissa, A. W., Narita, Y., "Vibrations of Completely Free Shallow Shells of Rectangular Planform," *Journal of Sound and Vibration*, Vol. 96, No. 2, pp. 207–218.

Lekhnitskii, S. G., 1968, *Anisotropic Plates*, (English translation), Gordon and Breach, New York.

Librescu, L., 1975, *Elastostatics and Kinetics of Anisotropic and Heterogeneous Shell-Type Structures*, Noordhoff, Loyden.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambridge University Press.

Lur'ye, A. I., 1940, "General Theory of Elastic Shells," (in Russian), *Prikl. Mat. Mekh.*, Vol. 4, No. 1, pp. 7–34.

Qatu, M. S., 1989, "Free Vibration and Static Analysis of Laminated Composite Shallow Shells," Ph.D. dissertation, Ohio State University.

Rao, K. P., 1978, "A Rectangular Laminated Anisotropic Shallow Thin Shell Finite Element," *Comp. Meth. Appl. Mech. Eng.*, Vol. 15, No. 1, pp. 13–33.

Vinson, J. R., and Sierakowski, R. L., 1986, *The Behavior of Structures Composed of Composite Materials*, Martinus Nijhoff, Dardrecht.

Whitney, J. M., 1987, *Structural Analysis of Laminated Anisotropic Plates*, Technomic Publishing, Inc., Lancaster, Penn.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

B. Yang
Assistant Professor,
Department of Mechanical Engineering,
University of Southern California,
Los Angeles, CA 90089-1453
Mem. ASME

C. D. Mote, Jr.
Professor and Chairman,
Department of Mechanical Engineering,
University of California, Berkeley,
Berkeley, CA 94720
Fellow ASME

Active Vibration Control of the Axially Moving String in the *S* Domain[1]

A new method is presented for active vibration control of the axially moving string, one of the most common models of axially moving continua. The control is formulated in the Laplace transform domain. The transfer function of a closed-loop system, consisting of the plant, a feedback control law and the dynamics of the sensing and actuation devices, is derived. Analysis of the root loci of the closed-loop system gives two stability criteria. Stabilizing controller design is carried out of both collocation and noncollocation of the sensor and actuator. It is found that all the modes of vibration can be stabilized and that in principle the spillover instability can be avoided. Also, the steady-state response of the stabilized string to periodic, external excitation is presented in closed form.

1 Introduction

Axially moving continua are the basic element in mechanical systems such as magnetic tape recorders, band saws, power transmission chains and belts, aerial cable tramways, and pipes containing flowing fluids (Mote, 1972). Vibration of the axially moving continua limits their utility in many applications, and in particular in high speed, precision systems. Vibration control is a key to improved performance in many of these systems.

Previous studies of vibration control focus on the introduction of passive dumping and stiffness (Wickert and Mote, 1988). Adaptation to complicated, time-varying conditions calls for vibration control through application of external, active forces. Axially moving continua belong to a class of distributed parameter systems. Control of distributed parameter systems is a research topic of increased interest (Balas, 1982; Johnson, 1982; Tzafestas and Stavroulakis, 1983). Active vibration control of axially moving continua, however, has developed slowly. Problems arise when one describes the axially moving continua with discrete models and uses state vector feedback in a control scheme. The uncontrolled or residual modes destabilize the system via spillover (Balas, 1978; Ulsoy, 1984). Also, due to the gyroscopic and distributed properties, construction of an observer to predict the motion of axially moving continua is difficult in both theory and practice. The dynamics of control actuators and sensors should be accounted for, which invariably complicates the control system formulation and controller design. From a practical point of view, due to operation constraints, for example, in the read/write process on a computer tape drive and in the wood cutting process on a band saw, the number of sensors and actuators used on axially moving continua is small.

In the present study a new method of active vibration control of axially moving continua is presented that addresses many of the difficulties summarized above. To demonstrate the development, the axially moving string is analyzed. The control objective is formulated in the Laplace transform domain. In the terminology of Butkovskiy (1983), the transfer function of a closed-loop system consisting of the plant, a feedback control law, and the dynamics of sensing and actuation devices is derived. Analysis of the root loci of the closed-loop system gives two stability criteria. Stabilizing controller design is carried out for both collocation and noncollocation of the sensor and actuator. It is found that all the modes of vibration can be stabilized and that the stability of the closed-loop system is presented in a polar plot of the controller. Because the physical measurement from the sensing device is used directly in the controller design, the effort needed to build an observer is eliminated. Finally, the steady-state response of the stabilized, closed-loop system to periodic, external excitation is presented in closed form.

II Problem Statement

A uniform axially moving string, traveling at a constant transport velocity between two fixed eyelets, is analyzed. The nondimensional equation of the transverse displacement $w(x,t)$ (Skutch, 1897) is given by

[1]This work was supported by the National Science Foundation, The University of California Forest Products Laboratory, and the author's group of industrial sponsors.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, Atlanta, Ga., Dec. 1-6, 1991.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 1988; final revision, Nov. 1989.

Paper No. 91-WA/APM-2.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$w_{tt}(x,t)+2cw_{xt}(x,t)-(1-c^2)w_{xx}(x,t)=f(x,t)\ \ x\epsilon[0,1]\ \ t\geq 0$$
$$w(0,t)=w(1,t)=0 \quad t\geq 0 \tag{1}$$

where $(\ )_t$ denotes $\frac{\partial}{\partial t}(\ )$, $(\ )_x$ denotes $\frac{\partial}{\partial x}(\ )$, c is the dimensionless transport speed, and $f(x,t)$ is the transverse external force. The initial conditions

$$w(x,t)\,|_{t=0}=u_0(x) \quad w_t(x,t)\,|_{t=0}=v_0(x) \tag{2}$$

are specified to complete the formulation. A critical velocity $c_{cr}=1$ exists at which the fundamental natural frequency vanishes and divergence occurs (Mote, 1965). In the present study assume that $c<1$.

In most cases point actuation and point sensing are more realizable than distributed actuation and sensing. According to Yang and Mote (1991) the axially moving string equation (1) is controllable and observable by a single point actuator and sensor which satisfy

$$\sin(k\pi x_a)\neq 0 \quad \sin(k\pi x_s)\neq 0 \quad x_a,x_s\epsilon(0,1) \quad \text{for } k=1,2,\ldots \tag{3}$$

where x_a and x_s are the locations of the actuator and the sensor, respectively. Satisfaction of equation (3) requires x_s and x_a to be irrational. With the above understanding, consider the vibration control of the axially moving string via one actuator and one sensor, i.e.,

$$w_{tt}(x,t)+2cw_{xt}(x,t)-(1-c^2)w_{xx}(x,t)=f(x,t)=\delta(x-x_a)u(t) \tag{4}$$

$$y(x)=\begin{cases} w(x_s,t) & \text{for displacement sensor} \\ w_t(x_s,t) & \text{for velocity sensor}\end{cases}$$

where $u(t)$ is the actuator output, and $y(t)$ is the sensor input.

The objective is to design a control law, which determines the actuator output $u(t)$ based on the measured information $y(t)$, such that the displacement of the string $w(x,t)$ is regulated asymptotically to zero for any initial disturbance.

III Transfer Function Formulation

The transfer function of a distributed system contains the information required to predict the system response under any initial and external disturbances (Butkovskiy, 1983). The control of a distributed system can be formulated in terms of transfer functions, and the analysis of control laws can be carried out in the s domain.

A Open-Loop Transfer Function. The Laplace transform of equation (1) with respect to t gives

$$s^2\bar{w}(x,s)+2cs\frac{\partial}{\partial x}\bar{w}(x,s)-(1-c^2)\frac{\partial^2}{\partial x^2}\bar{w}(x,s)=\bar{f}(x,s)+su_0(x)+v_0(x)+2c\frac{d}{dx}u_0(x) \tag{5}$$
$$\bar{w}(0,s)=\bar{w}(1,s)=0$$

where $\bar{w}(\cdot,s)$ and $\bar{f}(\cdot,s)$ are the Laplace transforms of $w(\cdot,t)$ and $f(\cdot,t)$ respectively, $u_0(x)$ and $v_0(x)$ represent the initial conditions (2), and s is a complex parameter. The transfer function of the axially moving string, $W_o(x,\xi,s)$, is the Green's function of the LHS of equation (5) (Butkovskiy, 1983). The solution of (5) is

$$\bar{w}(x,s)=\int_0^1 W_o(x,\xi,s)\left\{\bar{f}(\xi,s)+su_0(\xi)+v_0(\xi)+2c\frac{d}{d\xi}u_0(\xi)\right\}d\xi \tag{6}$$

where the transfer function $W_o(x,\xi,s)$ is given by

$$W_o(x,\xi,s)=\frac{K_o(x,\xi,s)}{D_o(s)} \tag{7a}$$

$$K_o(x,\xi,s)=\begin{cases}\frac{1}{s}\sinh(\gamma sx)\sinh(\gamma s(1-\xi))e^{\gamma cs(x-\xi)} & \text{for } 0\leq x\leq\xi\leq 1 \\ \frac{1}{s}\sinh(\gamma s\xi)\sinh(\gamma s(1-x))e^{\gamma cs(x-\xi)} & \text{for } 0\leq\xi\leq x\leq 1\end{cases} \tag{7b}$$

$$D_o(s)=\sinh(\gamma s) \quad \gamma=\frac{1}{1-c^2}. \tag{7c}$$

The roots of $D_o(s)$ from (7c),

$$p_k^o=ik\pi(1-c^2) \quad k=\pm 1,\pm 2,\ldots \quad i=\sqrt{-1},$$

are the eigenvalues of the axially moving string. Note that $s=0$ is not an eigenvalue because the limit of $W_o(x,\xi,s)$ as $s\to 0$ is determinate and finite.

B Closed-Loop Transfer Function. Now consider the control system described in equation (4). The control force (actuator) distribution is expressed in the transfer function form

$$\bar{f}(x,s)=\delta(x-x_a)G_a(s)G_c(s)G_s(s)\bar{w}(x_s,s) \tag{8}$$

where $G_a(s)$ and $G_s(s)$ are the transfer functions of the actuator at x_a and the sensor at x_s, respectively, and $G_c(s)$ represents the control law to be designed. The sensor, the control law, and the actuator are a feedback controller making a closed-loop system with the axially moving string. Define the transfer function of the feedback controller $C(s)$ by

$$C(s)\triangleq G_a(s)G_c(s)G_s(s). \tag{9}$$

A schematic of the closed-loop system is shown in Fig. 1.

Let the feedback controller have the form

$$C(s)=\mu\frac{K(s)}{D(s)}e^{-T_ds}=\mu\frac{k_0s^m+k_1s^{m-1}+\ldots+k_{m-1}s+k_m}{s^n+d_1s^{n-1}+\ldots+d_{n-1}s+d_n}e^{-T_ds}$$
$$\mu>0 \quad T_d\geq 0 \tag{10}$$

where the polynomials $K(s)$ and $D(s)$ have no common roots, T_d is the time delay, and μ is a gain parameter. Substitution of equations (8) and (9) into (6) gives

$$\bar{w}(x,s)=W_o(x,x_a,s)C(s)\bar{w}(x_s,s)+\int_0^1 W_o(x,\xi,s)\left\{su_0(\xi)+v_0(\xi)+2c\frac{d}{d\xi}u_0(\xi)\right\}d\xi. \tag{11}$$

Solve equation (11) for $\bar{w}(x_s,s)$ by letting $x=x_s$ and substitute $\bar{w}(x_s,s)$ in (11) to predict the closed-loop system response

$$\bar{w}(x,s)=\int_0^1 W_{cl}(x,\xi,s)\left\{su_0(\xi)+v_0(\xi)+2c\frac{d}{d\xi}u_0(\xi)\right\}d\xi \tag{12}$$

where the transfer function of the closed-loop system $W_{cl}(x,\xi,s)$ is obtained using equations (7) and (10)

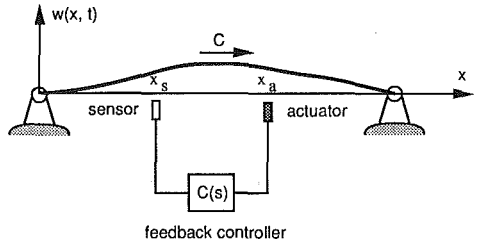


Fig. 1 Schematic of an axially moving string under feedback control

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$W_{cl}(x,\xi,s)=\frac{K_{cl}(x,\xi,s)}{D_{cl}(s)} \tag{13a}$$

$$K_{cl}(x,\xi,s)=K_o(x,\xi,s)D(s)+\mu\frac{K(s)}{D_o(s)}e^{-T_d s}[K_o(x_s,\xi,s)K_o(x,x_a,s)-K_o(x_s,x_a,s)K_o(x,\xi,s)] \tag{13b}$$

$$D_{cl}(s)=D_o(s)D(s)-\mu K_o(x_s,x_a,s)K(s)e^{-T_d s}. \tag{13c}$$

It is verified, by substituting equations 7(*b*), 7(*c*) into (13*b*), that the expression

$$\mu\frac{K(s)}{D_o(s)}e^{-T_d s}[K_o(x_s,\xi,s)K_o(x,x_a,s)-K_o(x_s,x_a,s)K_o(x,\xi,s)]$$

is analytic in s for all x and ξ. The eigenvalues of the closed-loop system (13) are the nonzero roots of the characteristic equation

$$D_{cl}(s)=D_o(s)D(s)-\mu K_o(x_s,x_a,s)K(s)e^{-T_d s}=0 \tag{14}$$

that depend on the parameters x_s, x_a, T_d, and μ.

For the system to be asymptotically stable all its eigenvalues must lie in the left-half s plane. The objective is to design a controller (10) so that the eigenvalues of the closed-loop system, determined from the roots of equation (14), all have negative real parts.

It should be noted that the transfer functions (7) and (13) are given in closed form; no approximation or truncation has been made. The high-frequency and low-frequency modes are treated equally in the formulation. Stability analyses address all modes of the system.

If for a certain root of $D_o(s)$, say p_k^o, $K_o(x_s,x_a,p_k^o)=0$, then by equation (14) $D_{cl}(p_k^o)=0$. This means that the feedback loop (8) has no influence on the kth mode and the system is not controllable and/or observable in that mode. In fact for $K_o(x_s,x_a,s_k^o)\neq 0$ for all k, the controllability and observability conditions given in equation (3) are recorded.

IV Stability Criteria

Let C_+ denote the open-half complex plane $\{s\epsilon C_+: \mathrm{Re}(s)>0\}$ and C_- the open-half complex plane $\{s\epsilon C_-: \mathrm{Re}(s)<0\}$. Let $\tilde{C}_+$ denote the closed-half complex plane $\{s\epsilon\tilde{C}_+: \mathrm{Re}(s)\geq 0\}$. Equation (13) describes an asymptotically stable system if all the eigenvalues of $D_{cl}(s)$ lie in C_-. If any eigenvalue lies in $\tilde{C}_+$, the system (13) is unstable.

Let the open-loop poles be the union of (i) the nonzero roots of $D_o(s)$, $p_k^o=ik\pi(1-c^2)$, (ii) the roots of $D(s)$, denoted by p_l, $l=1,2,\ldots n$, and (iii) the roots of $D_{cl}(s)$ at infinity in the complex plane C as $\mu\to 0^+$, denoted by p_q^∞. Let the open-loop zeros by the union of (i) the nonzero roots of $K_o(x_s,x_a,s)$, denoted by z_k^o, (ii) the roots of $K(s)$, denoted by z_l, $l=1,2,\ldots m$, and (iii) the roots of $D_{cl}(s)$ at infinity in the complex plane C as $\mu\to+\infty$, denoted by z_q^∞. The closed-loop eigenvalues trace continuous root loci in the complex plane starting at the open-loop poles and ending at the open-loop zeros as the parameter μ ranges from zero to infinity.

Consider the locations of the open-loop poles. First all the p_k^o are located on the imaginary axis. Because the controller (10) is stable, $p_l\epsilon C_-$, $l=1,2,\ldots n$. As shown in the Appendix, the p_q^∞ never fall in $\tilde{C}_+$ if $m-n\leq 1$, where m and n are the degrees of the polynomials $K(s)$ and $D(s)$, respectively. This condition is automatically satisfied in all usual cases. Consider that $p_q^\infty\epsilon C_-$. Accordingly, all the open-loop poles are in C_- except p_k^o.

Next examine the intersections of the root loci and the imaginary axis when $\mu>0$. Define

$$\theta\triangleq c(x_a-x_s)+T_d(1-c^2)$$

$$a\triangleq\min(x_s,x_a)\quad b\triangleq\max(x_s,x_a) \tag{15}$$

$$\frac{K(i\omega)}{D(i\omega)}\triangleq R(\omega)+iI(\omega)=M(\omega)e^{i\Phi(\omega)}$$

where $R(\omega)$ and $I(\omega)$ are real,

$$M(\omega)=\sqrt{R^2(\omega)+I^2(\omega)}\quad \Phi(\omega)=\tan^{-1}\left(\frac{I(\omega)}{R(\omega)}\right).$$

Using equations (7), (14), and (15) we obtain

$$\mu=\frac{\omega\sin(\gamma\omega)}{\sin(a\gamma\omega)\sin((1-b)\gamma\omega)M(\omega)}e^{i(\theta\gamma\omega-\Phi(\omega))}. \tag{16}$$

Let the intersections with the imaginary axis be $\pm i\alpha_k$, ordered

$$0<\alpha_1<\alpha_2<\ldots<\alpha_k<\ldots$$

Because μ is real and positive, α_k satisfy

$$\tan(\theta\gamma\alpha_k)=\tan(\Phi(\alpha_k))=\frac{I(\alpha_k)}{R(\alpha_k)}, \tag{17}$$

and the quantity

$$\mu_k\triangleq\frac{\alpha_k\sin(\gamma\alpha_k)}{\sin(a\gamma\alpha_k)\sin((1-b)\gamma\alpha_k)M(\alpha_k)}e^{i(\theta\gamma\alpha_k-\Phi(\alpha_k))} \tag{18}$$

is real and positive. For $\theta=0$ equation (17) gives $\frac{I(\alpha_k)}{R(\alpha_k)}=0$. Because $K(s)$ and $D(s)$ are polynomials in s, the number of solutions α_k and the number of intersections are finite. For $\theta\neq 0$, equation (17) has an infinite number of solutions α_k with $\alpha_k\to\infty$ as $k\to\infty$, and there is a corresponding infinite number of intersections.

Denote the root loci originating from p_k^o by $s_k(\mu)$. The direction of $s_k(\mu)$ at $\mu=0^+$ depends on $\left.\frac{ds_k}{d\mu}\right|_{\mu=0^+}$. By equations (7), (14), and (15) it is found that

$$\left.\frac{ds_k}{d\mu}\right|_{\mu=0^+}=\frac{K_o(x_s,x_a,p_k^o)K(p_k^o)e^{-T_d p_k^o}}{\left.\frac{d}{ds}D_o(s)\right|_{s=p_k^o}D(p_k^o)}=-\frac{i}{k\pi}\sin(k\pi x_s)\sin(k\pi x_a)\frac{K(p_k^o)}{D(p_k^o)}e^{-ik\pi\theta}. \tag{19}$$

If for all k $\mathrm{Re}\left(\left.\frac{ds_k}{d\mu}\right|_{\mu=0^+}\right)<0$, all $s_k(\mu)$ move into C_-. Further let μ increase continuously from zero until the first root locus intersects the imaginary axis while passing into C_+ at some critical value μ_{cr}. This critical value, μ_{cr}, is

$$\mu_{cr}=\inf_k\{\mu_k\} \tag{20}$$

where μ_k is given in equation (18). In the case of no intersection between any root locus and the imaginary axis, μ_{cr} is defined to be infinite. For $\mu\epsilon(0,\mu_{cr})$, all the closed-loop eigenvalues belong to C_-, and as a consequence the control system (13) is stable.

If $\mathrm{Re}\left(\left.\frac{ds_k}{d\mu}\right|_{\mu=0^+}\right)>0$ holds for an infinite number of root loci, those root loci enter C_+ as μ increases from zero. When $\theta=0$, as mentioned earlier, a finite number of loci intersect the imaginary axis. In this case there is always an infinite number of root loci in C_+ for any $\mu>0$. When $\theta\neq 0$ there are infinite intersection points $\pm i\alpha_k$ with $\alpha_k\to\infty$ as $k\to\infty$. Let $\eta_k(\mu)$ be the root locus that intersects the imaginary axis at $i\alpha_k$. That is, $\eta_k(\mu_k)=i\alpha_k$, where μ_k is given in equation (18). The

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$\mathrm{Re}\left(\frac{d\eta_k}{d\mu}\Big|_{\mu=\mu_k}\right)$ determines if η_k enters or leaves C_+. By equations (7), (10), and (14) it can be shown that as $k\to\infty$

$$\mathrm{Re}\left(\frac{d\eta_k}{d\mu}\Big|_{\mu=\mu_k}\right)\to\frac{1}{\gamma\mu_k}\frac{\theta}{[\cot(\gamma\alpha_k)-a\cot(a\gamma\alpha_k)-(1-b)\cot((1-b)\gamma\alpha_k)]^2+\theta^2}.$$

This implies that there exists an integer $\tilde{K}>0$ such that for $k>\tilde{K}$, the root locus $\eta_k(\mu)$ always enters (leaves) C_+ if $\theta>0$ (<0). If $\theta>0$ there are at most $\tilde{K}$ $\eta_k(\mu)$ leaving C_+. All other root loci which are in C_+ initially will stay in C_+ as μ increases. If $\theta<0$ at most $\tilde{K}$ $\eta_k(\mu)$ enter C_+. All other (infinite) root loci $\eta_k(\mu)$ leave C_+. As shown in the Appendix, for $\theta\neq 0$ there is an infinite number of open-loop zeros z_k^∞ in C_+. Because every root locus will terminate at an open-zero, those $s_k(\mu)$ which enter C_+ initially cannot all intersect the imaginary axis and remain in C_-. Otherwise some open-loop z_k^∞ would not be the end points of root loci, which is a contradiction. So there are always root loci in C_+ for any $\mu>0$. Thus, the system (13) is always unstable if $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)>0$ holds for an infinite number of root loci.

For $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)=0$, the stability of the closed-loop system (13) cannot be determined directly. If $\mathrm{Im}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)$ is also zero, equation (19) gives $\sin(k\pi x_s)\sin(k\pi x_a)\frac{K(p_k^o)}{D(p_k^o)}=0$. This indicates that the system is uncontrollable or unobservable for the kth mode. The root locus s_k is the point p_k^o. If $\mathrm{Im}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)$ is not zero, further investigation for higher terms, at least $\mathrm{Re}\left(\frac{d^2s_k}{d\mu^2}\Big|_{\mu=0^+}\right)$, is needed to predict local stability.

In summary we give the following two criteria.

Criterion 1. The control system (13) is stable for $\mu\epsilon(0,\mu_{cr})$ if $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)<0$ for all k.

Criterion 2. The control system (13) is unstable for any $\mu>0$ if $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)>0$ for an infinite number of k.

In controller design, stability analysis of the closed-loop system requires the following two steps: (1) find a control law (14) such that $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)<0$ for all k and (2) determine the stable range $(0,\mu_{cr})$ for μ.

V Design of Particular Controllers

Based on the results in last section, the problem of stabilizing the displacement of the axially moving string by the feedback controller (10) is studied in three cases. Of the major concern is spillover, the destabilizing effect of residual modes resulting from the observation/control of a finite number of modes (Balas 1978).

A Collocation of Sensor and Actuator ($x_s=x_a$). Let $x_s=x_a=d$, and consider feedback control (10) without time delay, i.e., $T_d=0$. Then $\theta=0$, and equation (19) becomes

$$\frac{ds_k}{d\mu}\Big|_{\mu=0^+}=-\frac{i}{k\pi}\sin^2(k\pi d)\frac{K(p_k^o)}{D(p_k^o)}\quad p_k^o=ik\pi(1-c^2). \tag{21}$$

First, we investigate the stability of the axially moving string under negative velocity feedback

$$C(s)=-\mu s\quad K(s)=-s\quad D(s)=1. \tag{22}$$

Substituting equation (22) into (21) gives $\frac{ds_k}{d\mu}\Big|_{\mu=0^+}=-(1-c^2)\sin^2(k\pi d)$. Because the sensor and actuator are placed such that the string is controllable and observable for all modes, $\sin(k\pi d)\neq 0$ for $k=1,2,\ldots$ by equation (3). This renders $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)<0$ for all k. Also, equation (17) reduces to $\alpha=0$, implying $\mu_{cr}\to\infty$. According to Criterion 1, the axially moving string under the velocity feedback equation (22) is stable for any positive μ.

For general $\frac{K(s)}{D(s)}$, rewrite equation (21) using (15)

$$\frac{ds_k}{d\mu}\Big|_{\mu=0^+}=\frac{1}{k\pi}\sin^2(k\pi d)M(\omega_k^o)e^{i\left(\Phi(\omega_k^o)-\frac{\pi}{2}\right)}\quad \omega_k^o=k\pi(1-c^2). \tag{23}$$

It is seen that $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)<0$ if and only if $\Phi(\omega_k^o)\epsilon(\pi,2\pi)$. However, $\Phi(\omega_k^o)\epsilon(\pi,2\pi)$ is equivalent to $I(\omega_k^o)<0$. Hence, $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)<0$ for all k if $I(\omega)<0$ for any $\omega>0$. Next, determine the parameter μ_{cr}. Substitution of $\theta=0$ into equation (17) gives $I(\alpha)=0$. Assume $I(\omega)<0$, $\omega\epsilon(0,\infty)$. Then the above equation has no positive roots, which implies that $\mu_{cr}\to\infty$. Thus, by Criterion 1 we conclude:

Claim 1. The axially moving string, under the feedback control law $C(s)=\mu\frac{K(s)}{D(s)}$ with collocated sensor and actuator, is stable for any $\mu>0$ if for $\omega\epsilon(0,\infty)$

$$I(\omega)=\mathrm{Im}\left(\frac{K(i\omega)}{D(i\omega)}\right)<0.$$

Claim 1 requires that the polar plot $\frac{K(i\omega)}{D(i\omega)}$ lie below the real axis. This condition can be satisfied by many transfer functions. Here are some examples of stabilizing controllers:

(a) $\frac{K(s)}{D(s)}=\frac{1+T_1s}{1+T_2s}$, $0<T_1<T_2$: $I(\omega)=\frac{(T_1-T_2)\omega}{1+\omega^2T_2^2}<0$ for $\omega>0$.

(b) $\frac{K(s)}{D(s)}=\frac{-s}{1+T_s}$, $T>0$: $I(\omega)=\frac{-\omega}{1+\omega^2T^2}<0$ for $\omega>0$.

(c) $\frac{K(s)}{D(s)}=\frac{1}{(1+T_1s)(1+T_2s)}$, T_1 and $T_2>0$: $I(\omega)=\frac{-(T_1-T_2)\omega}{1+\omega^2T_1^2)(1+\omega^2T_2^2)}<0$ for $\omega>0$.

In short, with collocation of the sensor and actuator all the modes in vibration can be stabilized, and spillover instability can be avoided.

B Noncollocation of Sensor and Actuator Without Time Delay ($x_s\neq x_a$, $T_d=0$). Noncollocation of sensors and actuators can cause spillover. To show this consider the velocity feedback (22) with $x_s\neq x_a$. In this case equation (19) reduces to

$$\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)=-(1-c^2)\sin(k\pi x_s)\sin(k\pi x_a)\cos(k\pi c(x_s-x_a)).$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

For any x_a, $x_s \epsilon (0,1)$, $x_s \neq x_a$, there always exists an infinite number of k such that $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right) > 0$. According to Criterion 2 the axially moving string under the velocity feedback (22) is always unstable. It is interesting to see that for any integer $N>0$, if x_s and x_a are sufficiently close together, $\sin(k_l \pi x_s)\sin(k_l \pi x_a)\cos(k_l \pi c(x_s - x_a)) > 0$ for $k = 1,2,\ldots N$. With properly chosen μ the eigenvalues $s_k(\mu)$, $k = 1,2,\ldots N$, will lie in C_-. The N modes are stabilized. But with $x_a \neq x_s$ an infinite number of modes are destabilized by the controller (22). As $x_a - x_s$ approaches zero, an increasing of modes are stabilized, until we return to the case of collocated sensor and actuator.

In general, any feedback controller (10) with $T_d = 0$ and noncollocation of the sensor and actuator leads to an infinite number of unstable modes. The sequence $\{\sin(k\pi x_s)\sin(k\pi x_a)e^{-ik_\pi c(x_a - x_s)}\}$ $k = 1,2,\ldots$, has an infinite number of points in each of the four quadrants of the complex plane. As $k \to \infty$, $\frac{K(p_k^o)}{D(p_k^o)} \to k_0 (ik\pi(1-c^2))^{m-n}$, where m and n are the orders of $K(s)$ and $D(s)$, respectively, which belongs to one of the four quadrants of the complex plane. By equation (19) the derivative $\frac{ds_k}{d\mu}\Big|_{\mu=0^+}$ will have positive real part for an infinite number of k. According to Criterion 2 the control system (13) is unstable for any $\mu > 0$.

C Noncollocation of Sensor and Actuator With Time Delay ($x_s \neq x_a$, $T_d \neq 0$). The feedback controller (10) with a specific time delay T_d and the symmetric positioning of the sensor and actuator can stabilize the modes. Let the sensor and actuator be positioned at the points that are symmetric with respect to the midpoint of the string, i.e., $x_s + x_a = 1$. Equation (19) becomes

$$\frac{ds_k}{d\mu}\Big|_{\mu=0^+} = = (-1)^k \frac{i}{k\pi}\sin^2(k\pi x_s)\frac{K(p_k^o)}{D(p_k^o)}e^{-ik\pi\theta}.$$

For $\theta = 1$ ($T_d = \frac{1}{1-c^2}[1 + c(x_s - x_a)]$ in (15)), the above becomes

$$\frac{ds_k}{d\mu}\Big|_{\mu=0^+} = \frac{1}{k\pi}\sin^2(k\pi x_s) M(\omega_k^o) e^{i\left(\Phi(\omega_k^o) + \frac{\pi}{2}\right)}$$

$$\omega_k^o = k\pi(1-c^2). \tag{24}$$

It is seen that $\mathrm{Re}\left(\frac{ds_k}{d\mu}\Big|_{\mu=0^+}\right) < 0$ for all k if $I(\omega) > 0$ for any $\omega > 0$. Assume $I(\omega)$ is positive for $\omega > 0$. By $\theta = 1$, $x_s + x_a = 1$ and (18), the parameter μ_{cr} is found to be

$$\mu_{cr} = \inf_k \frac{\alpha_k I(\alpha_k)}{M^2(\alpha_k)\sin^2(a\gamma\alpha_k)} \quad a = \min(x_s, x_a) \tag{25}$$

where α_k satisfies

$$\tan(\gamma\alpha_k) = \frac{I(\alpha_k)}{R(\alpha_k)} \quad k = 1,2,\ldots \tag{26}$$

Applying Criterion 1 we obtain:

Claim 2. The axially moving string, under the feedback controller (10) with noncollated sensor and actuator, is stable if

$$I(\omega) = \mathrm{Im}\left(\frac{K(i\omega)}{D(i\omega)}\right) > 0 \quad \omega \epsilon (0,\infty) \tag{27}$$

$$x_s + x_a = 1 \quad T_d = \frac{1}{1-c^2}[1 + c(x_s - x_a)] \quad \mu \epsilon (0, \mu_{cr})$$

where μ_{cr} is given in equation (25).

A physical explanation of this control highlights special properties of the axially moving string. The kth mode in free vibration is described by

$$w_k(x,t) = A_k \sin(k\pi x)\sin(k\pi(1-c^2)t + k\pi c x + \phi_k) \tag{28}$$

where A_k and ϕ_k are real constants. With x_s, x_a, and T_d satisfying equation (27),

$$w_k(x_a,t) = -w_k(x_s, t - T_d) \quad \frac{\partial}{\partial t} w_k(x_a,t) = -\frac{\partial}{\partial t} w_k(x_s, t - T_d). \tag{29}$$

These state that the transverse motion of the string at $x = x_a$ is exactly predicted by the displacement at $x = x_s$ at time T_d earlier. A control force of proper magnitude and phase can damp the kth mode. Because T_d is identical for all modes, the control force suppresses all modes simultaneously. The parameter μ is not arbitrarily large for the noncollocated sensor and actuator. The string displacement will be destabilized if $\mu > \mu_{cr}$.

Claim 2 requires that the polar plot $\frac{K(i\omega)}{D(i\omega)}$ lie above the real axis for stability. Recall that for $x_s = x_a$ stability requires the polar plot $\frac{K(i\omega)}{D(i\omega)}$ lie below the real axis. All the stabilizing controllers satisfying Claim 1 then can be used here by adding a phase shift of 180 deg. As an example, instead of negative velocity feedback (22), positive velocity feedback

$$C(s) = \mu s e^{-T_d s} \quad K(s) = s \quad D(s) = 1, \tag{30}$$

where x_s, x_a, and T_d satisfy equation (27), is a stabilizing control. In this case

$$I(\omega) = \omega \quad \mu_{cr} = \inf_k \frac{1}{\sin^2\left(a\left(2k + \frac{1}{2}\right)\pi\right)} = 1.$$

Since $I(\omega)$ is positive for $\omega \epsilon (0,\infty)$, by Claim 2, the axially moving string under the positive feedback controller (30) is stable for $\mu \epsilon (0,1)$.

VI Frequency Response

Once stability is verified, the steady-state response of the axially moving string to periodic, external excitations can be analyzed. Let the force distribution in equation (5) be

$$\bar{f}(x,s) = \bar{f}_d(x,s) + \delta(x - x_a)C(s)\bar{w}(x_s,s)$$

where $\bar{f}_d(x,s)$ is the Laplace transform of $f_d(x,t)$. The response of the closed-loop system, which is equation (12) with the above equation replacing (8), is

$$\bar{w}(x,s) = \int_0^1 W_{cl}(x,\xi,s)\Big\{\bar{f}_d(\xi,s) + s u_0(\xi) + v_0(\xi) + 2c\frac{d}{d\xi}u_0(\xi)\Big\}d\xi. \tag{31}$$

First consider sinusoidal excitation

$$f_d(x,t) = q(x)e^{i\omega t} \quad \bar{f}_d(x,s) = q(x)\frac{1}{s - i\omega}. \tag{32}$$

Substitution of equation (32) and (13*a*) into (31) yields

$$\bar{w}(x,s) = \frac{M_{cl}(x,s)}{(s - i\omega)D_{cl}(s)} + \frac{N_{cl}(x,s)}{D_{cl}(s)}$$

where

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

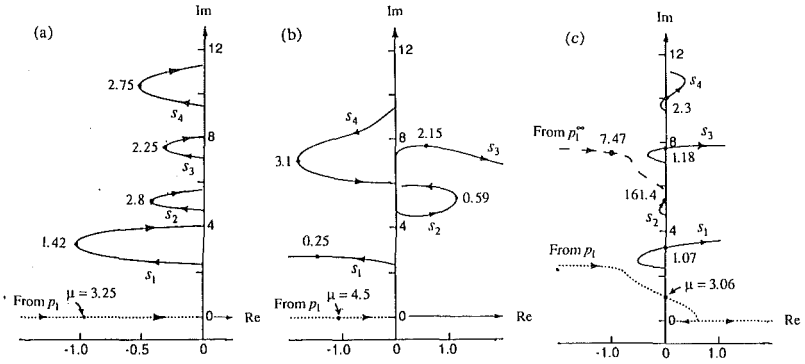


Fig. 2 Root loci of the string: (*a*) under the controller (38) (collocation of the sensor and actuator); (*b*) under the controller (39) (dislocation of the sensor and actuator with zero time delay); (*c*) under the controller (40) (dislocation of the sensor and actuator with time delay)

$$M_{cl}(x,s) = \int_0^1 K_{cl}(x,\xi,s)q(\xi)d\xi$$

$$N_{cl}(x,s) = \int_0^1 K_{cl}(x,\xi,s)\left\{su_0(\xi) + v_0(\xi) + 2c\frac{d}{d\xi}u_0(\xi)\right\}d\xi.$$

Because $M_{cl}(x,s)$, $N_{cl}(x,s)$ and $D_{cl}(s)$ are analytic functions in s, the inverse Laplace transform of $\bar{w}(x,s)$ leads to

$$w(x,t) = \sum_{k=1}^{\infty} g_k(x)e^{\lambda_k t} + h(x)e^{i\omega t}$$

where $\lambda_k(\mathrm{Re}(\lambda_k)<0)$ are the closed-loop eigenvalues determined from $D_{cl}(s)=0$, and

$$g_k(x) = \operatorname*{Res}_{s=\lambda_k}\left(\frac{N_{cl}(x,s)}{D_{cl}(s)}\right) + \operatorname*{Res}_{s=\lambda_k}\left(\frac{M_{cl}(x,s)}{(s-i\omega)D_{cl}(s)}\right)$$

$$h(x) = \operatorname*{Res}_{s=i\omega}\left(\frac{M_{cl}(x,s)}{(s-i\omega)D_{cl}(s)}\right) = \frac{M_{cl}(x,i\omega)}{D_{cl}(i\omega)}.$$

The steady-state response of the string is

$$w(x,t) = h(x)e^{i\omega t} = \int_0^1 W_{cl}(x,\xi,i\omega)q\xi d\xi e^{i\omega t}. \tag{33}$$

For a harmonic excitation is at a point x_f

$$f_d(x,t) = q_0\delta(x-x_f)e^{i\omega t} \quad x_f\epsilon(0,1), \tag{34}$$

the steady-state response is

$$w(x,t) = W_{cl}(x,x_f,i\omega)q_0e^{i\omega t}. \tag{35}$$

For an excitation $f(x,t)$ of period T

$$f_d(x,t) = \sum_k q_k(x)e^{i\omega_k t} \quad \omega_k = \frac{2\pi k}{T}, \tag{36}$$

by superposition, the steady-state response is

$$w(x,t) = \sum_k \int_0^1 W_{cl}(x,\xi,i\omega_k)q_k(\xi)d\xi e^{i\omega_k t}. \tag{37}$$

Thus, with the knowledge of $W_{cl}(x,\xi,s)$ the steady-state response of the axially moving string under periodic excitations is completely determined. Note that the steady-state responses (33) and (37) are given in closed form.

VII Numerical Examples

The theory is illustrated on an axially moving string at transport velocity $c=0.5$ under velocity feedback. The actuator is described as a first-order system $G_a(s) = \dfrac{k_a}{1+T_a s}$, where k_a is an adjustable gain, and T_a is the time constant. Let $T_a=0.1$. The actuator then has a pole at $p_1 = \dfrac{-1}{T_a} = -10.0$. Assume a displacement sensor with a negligible time constant is used. So $G_s(s)=k_s$, where k_s is a constant gain. Three cases are considered.

Table 1

Mode number k	$\mathrm{Re}\left(\dfrac{ds_k}{d\mu}\Big\|_{\mu=0^+}\right)$
1	−0.6368
2	0.2089
3	0.1186
4	−0.1748

Case A. Collocation of the sensor and actuator

$$G_c(s) = -s \quad C(s) = \frac{-\mu s}{1+0.1s}$$

$$x_s = x_a = 0.4167 \quad \mu = k_a k_a. \tag{38}$$

Here x_a and x_s are the irrational numbers with four significant digits. According to Claim 1, for any $\mu>0$, the controller (38) stabilizes the axially moving string. The root loci of the first four modes of the string (in solid lines) are shown in Fig. 2(*a*). The dotted line in the figure is the root locus starting from the actuator pole p_1. As $\mu\to\infty$, the root loci converge to the imaginary axis.

Case B. Noncollocation of the sensor and actuator without time delay

$$G_c(s) = -s \quad C(s) = \frac{-\mu s}{1+T_a S}$$

$$x_s = 0.4167 \quad x_a = 0.6233 \quad \mu = k_a k_s. \tag{39}$$

In this case $\mathrm{Re}\left(\dfrac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)$ for the first four modes are calculated using (19) and given in Table 1. Because the $\mathrm{Re}\left(\dfrac{ds_k}{d\mu}\Big|_{\mu=0^+}\right)$ are positive for $k=2,3$ and negative for $k=1,4$, the second and third modes are unstable, and the first and fourth modes are stable at $\mu=0^+$. The axially moving string under equation (39) is always unstable for any given $\mu>0$. Shown in Fig. 2(*b*) are the root loci of the string under the controller (39). The dotted line is the root locus starting from the actuator pole p_1.

Case C. Noncollocation of the sensor and actuator with time delay

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

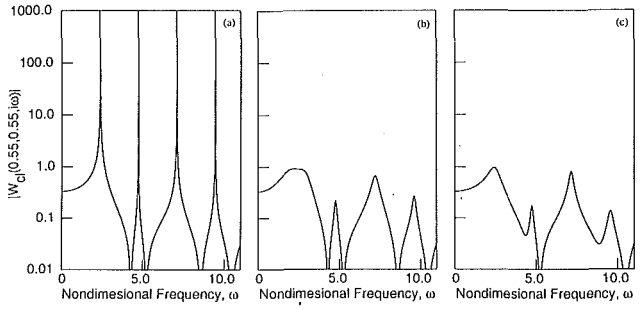


Fig. 3 Receptance of the string: (*a*) without control; (*b*) under the controller (38) (collocation of the sensor and actuator); (*c*) under the controller (40) (dislocation of the sensor and actuator with time delay)

$$G_c(s)=s \quad C(s)=\frac{\mu s}{1+T_a s}e^{-T_d s}$$

$$x_s=0.4167 \quad x_a=1-x_s=0.5833 \quad \mu=k_a k_s$$

$$T_d=\frac{1}{1-c^2}[1+c(x_s-x_a)]=1.2223. \tag{40}$$

Following Section V.A, example (*b*), $I(\omega)>0$ for all $\omega>0$. By equation (25), $\mu_{cr}=1$. According to Claim 2, the axially moving string is stable for $\mu\epsilon(0,1)$. The root loci of the control system are drawn in Fig. 2(*c*). The second and third modes, which are unstable in Case B, are stabilized. The dashed locus originates at infinity because $e^{-\theta s}$ in the characteristic equation (14) introduces an infinite number of open-loop poles $p_m^\infty \epsilon C_-$ (see the Appendix). This also occurs in Case B where the factor $e^{\gamma c s(x_s-x_a)}$ appears in equation (14). For the purpose of concentrating on the unstable modes, the root loci beginning from p_m^∞ are not shown in Fig. 2(*b*).

Next, consider the frequency response of the stabilized string to the harmonic excitation (34) at point $x_f=0.55$. The response is calculated at the point of excitation $x=0.55$. The receptance is drawn for the cases of the string without and with control in Fig. 3. The active controls (38) and (40) reduce the resonance peak amplitudes.

Finally, the steady-state response of the axially moving string to a periodic impulse (Fig. 4(*a*)) at $x_f=0.55$ is evaluated using equation (37). Shown in Fig. 4(*b*), (*c*), and (*d*) are the responses of the string, at $x=0.55$, without control and with the controls (38) and (40), respectively.

VIII Conclusion

An active vibration control of axially moving continua has been presented here. Spillover instability can be avoided for both collocation and noncollocation of the sensor and actuator. With the transfer function formulation, the observer is eliminated and the dynamics of the sensor and actuator are accounted for without complicating the controller design. The steady-state responses of the stabilized systems are evaluated in closed form. Although this study has focused on axially moving continua, the transfer function concept and the time delay approach can be utilized in vibration control of other flexible systems. In addition, the method presented readily extends to systems with multiple sensors and actuators.

References

Balas, M. J., 1978, "Feedback Control of Flexible System," *IEEE Automatic Control*, Vol. AC-23, No. 4, pp. 673–679.

Balas, M. J., 1982, "Trends in Large Space Structure Control Theory: Fondest Hopes, Widest Dreams," *IEEE Automatic Control*, Vol. AC-27, No. 3, pp. 15–33.

Butkovskiy, A. G., 1983, *Structural Theory of Distributed Systems*, Ellis Horwood Limited.

Johnson, T. L., 1982, "Progress in Modeling and Control of Flexible Spacecraft," *Journal of The Franklin Institute*, Vol. 315, pp. 495–520.

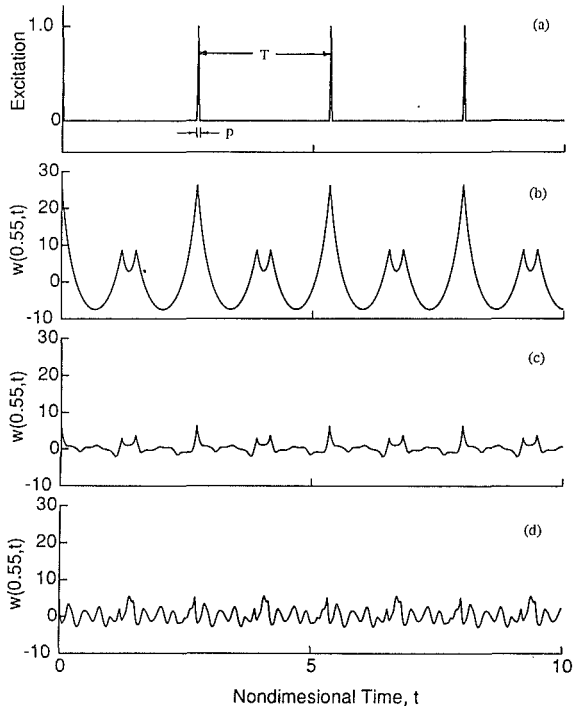


Fig. 4 Steady-state response of the string under a periodic impulse: (*a*) excitation: $T=2.37$, $p=0.05$; (*b*) response of the string without control; (*c*) response of the string with the controller (38); (*d*) response of the string with the controller (40)

Mote, C. D., Jr., 1965, "A Study of Band Saw Vibrations," *Journal of The Franklin Institute*, Vol. 279, No. 6, pp. 430–444.

Mote, C. D., Jr., 1972, "Dynamic Stability of Axially Moving Materials," *Shock and Vibration Digest*, Vol. 4, No. 4, pp. 2–11.

Skutch, R., 1897, "Uber die Bewegung Eines Gespannten Fadens, Weicher Gezwungun ist, Durch Zwei Feste Punkte, mit Einer Constantan Geschwindigkeit zu gehen, und Zwischen denselben in Transversal Schwingungen von gerlinger Amplitude Versetzi Wird," *Annalen der Physik Chemie*, Vol. 61, pp. 190–195.

Tzafestas, S. G., and Stavroulakis, P., 1983, "Recent Advances in the Study of Distributed Parameter Systems," *Journal of The Franklin Institute*, Vol. 315, No. 5-6, pp. 285–305.

Ulsoy, A. G., "Vibration Control in Rotating or Translating Elastic Systems," *ASME Journal of Dynamic Systems, Measurement, and Control*, Vol. 106, No. 1, pp. 6–14.

Wickert, J. A., and Mote, C. D., Jr., 1988, "Current Research on the Vibration and Stability of Axially Moving Materials," *Shock and Vibration Digest*, Vol. 20, No. 5, pp. 3–13.

Yang, B., and Mote, C. D., Jr., "On Controllability and Observability Distributed Gyroscopic Systems," to *ASME Journal of Dynamic Systems, Measurement, and Control*, in press.

APPENDIX

Locations of the Open-Loop Poles p_q^∞ and Open-Loop Zeros z_q^∞

The characteristic equation (14) can be written as

$$\frac{1}{\mu}=f(s)\equiv\frac{K(s)}{sD(s)}\frac{\sinh(a\gamma s)\sinh((1-b)\gamma s)}{\sinh(\gamma s)}e^{-\theta\gamma s}, \tag{A1}$$

where $\mu\epsilon(0,\infty)$, and a, b, and θ are given in (15). According to the definitions of p_q^∞ and z_q^∞,

$$f(s)\to+\infty \quad \text{as } s\to p_q^\infty,\ |s|\to\infty \tag{A2a}$$

$$f(s)\to 0^+ \quad \text{as } s\to z_q^\infty,\ |s|\to\infty. \tag{A2b}$$

Write $s=x+iy$, where $x,y\epsilon R$ and $i=\sqrt{-1}$. There are three possibilities for $|s|\to\infty$: (i) $x\to\pm\infty$, $y\to$a finite number; (ii) $y\to\pm\infty$, $x\to$a finite number, and (iii) x, $y\to\pm\infty$. Only the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

case (i) can render $f(s)\to\infty$ or 0. So, to determine p_q^∞ and z_q^∞ one only needs to consider the case when $s\to\pm\infty+iy$, $|y|<\infty$. By $\dfrac{K(s)}{D(s)}$ in equation (10),

$$f(s)\to\frac{1}{2}k_0(x+iy)^{m-n-1}e^{-(x+iy)\psi}\quad\text{as } s\to+\infty+iy \tag{A3a}$$

$$f(s)\to-\frac{1}{2}k_0(x+iy)^{m-n-1}e^{-(x+iy)\phi}\quad\text{as } s\to-\infty+iy \tag{A3b}$$

where

$$\psi=\gamma(\theta+b-a)\quad\phi=\gamma(\theta+a-b). \tag{A3c}$$

Note that ψ is non-negative by equation (15).

First, consider $m-n-1=0$. For $\psi=0$ ($\phi=0$) equation (A3*a*), (A3*b*) yields $f(s)\to\dfrac{1}{2}k_0\left(f(s)\to-\dfrac{1}{2}k_0\right)$. There is neither p_q^∞ nor z_q^∞. For $\psi>0$, $f(s)\to0$ as $x\to+\infty$ equation (A3*a*). There is only an infinite number of open-loop zeros in C_+.

$$z_q^\infty=+\infty+i\frac{\pi}{\psi}\left[2q+\frac{1}{2}(1-\operatorname{sgn}(k_0))\right]$$

$$q=0,\pm1,\pm2,\ldots \tag{A4}$$

For $\phi>0$ ($\phi<0$) $f(s)\to\infty$ ($f(s)\to0$) as $x\to-\infty$ (A3*b*). There is only an infinite number of open-loop poles (zeros in C_-

$$p_q^\infty(z_q^\infty)=-\infty+i\frac{\pi}{\phi}\left[2q+\frac{1}{2}(1+\operatorname{sgn}(k_0))\right]$$

$$q=0,\pm1,\pm2,\ldots \tag{A5}$$

Next, consider $m-n-1<0$. For $\psi=0$ $f(s)\to0$ as $x\to+\infty$ (A3*a*). There is at most one open-loop zero in C_+

$$z_1^\infty=+\infty\quad\text{if } k_0>0. \tag{A6}$$

For $\phi=0$ $f(s)\to0$ as $x\to-\infty$ (A3*b*). There is at most one open-loop zero in C_-

$$z_1^\infty=-\infty\quad\text{if } (-1)^{m-n}k_0>0. \tag{A7}$$

For $\psi>0$ $f(s)\to0$ as $x\to\infty$. There is only an infinite number of open loop poles zeros in C_+

$$z_q^\infty=+\infty+i\frac{\pi}{\psi}\left[2q+\frac{1}{2}(1-\operatorname{sgn}(k_0))\right]$$

$$q=0,\pm1,\pm2,\ldots \tag{A8}$$

For $\phi>0$ ($\phi<0$) $f(s)\to\infty$ ($f(s)\to0$) as $x\to-\infty$. There is only an infinite number of open-loop poles (zeros) in C_-

$$p_q^\infty(z_q^\infty)=-\infty+i\frac{\pi}{\phi}\left\{2q+\frac{1}{2}[1-\operatorname{sgn}((-1)^{m-n}k_0)]\right\}$$

$$q=0,\pm1,\pm2,\ldots \tag{A9}$$

In summary, the open-loop poles p_q^∞ never fall in $\tilde{C}_+$ for $m-n\le1$. Futhermore, if $\theta\neq0$ (so $\psi>0$), there is an infinite number of open-loop zeros z_q^∞ in C_+.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Y.-A. Su

I. G. Tadjbakhsh
Fellow ASME

Department of Civil Engineering,
Rensselaer Polytechnic Institute,
Troy, NY 12180-3590

Optimal Control of Beams With Dynamic Loading and Buckling

The vibration of beams can be suppressed by a set of discrete actuators associated with a causal optimal, coupled-modal control algorithm which is derived from minimizing a time-dependent weighted sum of the kinetic energy, potential energy, and the control effort. The control gains are functions of the control interval and couple the modes that participate in the control. The controller can be applied to control a beam subjected to a resonant cyclic load, a moving load, and a beam column subjected to a unit impuse at midspan before buckling.

Introduction

Active control of Euler-Bernoulli beam can be formulated in terms of spatial modal functions of the governing self-adjoint partial differential equation and the boundary conditions. The control is achieved by controlling the modal amplitudes. This is the basic idea of modal control of distributed-parameter systems. The modal control system can be designed by either pole allocation technique (Simon and Mitter, 1968) or by the standard optimal control using a quadratic performance measure (Kirk, 1970). Meirovitch et al. (1980, 1982, 1983, 1985) have developed the method of independent modal space control (IMSC). In this method each mode of the system is considered to be controllable independently.

The problem of active control of a distributed-parameter system with discrete sets of sensors and actuators is of the category of coupled modal control (Meirovitch, 1987). Here, an optimal, coupled-modal control algorithm is developed for continuous self-adjoint structures by minimizing a time-dependent weighted sum of the kinetic energy, potential energy, and the control effort. This leads to a causal optimal algorithm whereby control forces are determined solely on the basis of information available up to the time at which control is being implemented. Its control gains are therefore functions of control interval and couple with each of the participated control modes (Tadjbakhsh and Su, 1989).

The method is used to suppress the vibration of an undamped beam subjected to a periodic load resonant with the fundamental frequency of the beam. The control of the beam under a moving load is also considered. Both cases show that the control interval affects the behavior of the controlled system significantly. The control algorithm is applied to a beam column subjected to an impulse load at midspan. The vibrations are suppressed when the axial load is below the first buckling load of the column.

Control Problem

Consider an undamped beam with a length l. Let $z(x, t)$ denote the deflection of the beam with x and t, respectively, representing the position of a point on the beam and the time. Assume n actuators at location x_i, $i=1, 2, \ldots n$ with control forces u_i. The equation of motion of the beam can be written in the form.

$$m(x)\ddot{z}+Lz=f(x, t)+\sum_{i=1}^{n} u_i(t)\delta(x-x_i),\ t>t_0,\ 0<x<l \quad (1)$$

where $m(x)$ is the mass distribution, $f(x, t)$ is the external excitation, δ is the Dirac delta function, and L is a spatial operator expressing the stiffness of the beam and can be given by

$$L=\frac{\partial^2}{\partial x^2}\left[EI(x)\frac{\partial^2}{\partial x^2}\right], \quad (2)$$

where E is Young's modulus of the material and $I(x)$ is the second moment inertia of the cross-section of the beam. In equation (1), the superimposed dots denote time derivatives. The displacement $z(x, t)$ is subject to the boundary conditions of the type

$$B_i z=0, \quad i=1,2 \quad (3)$$

at each boundary point $x=0$ and $x=l$ and B_is are linear differential operators of order ranging from zero to three. Inhomogeneous boundary cconditions can be accommodated by using appropriate transformations that modify z and f and reduce the problem to the homogeneous boundary conditions.

The initial time t_0 is any arbitrary point on the time scale at which the initial conditions are

$$z_0(x)=z(x, t_0),\ \dot{z}_0(x)=\dot{z}(x, t_0). \quad (4)$$

The associated eigenvalue problem consists of the differential equations

$$L\phi_j=\omega_j^2 m\phi_j,\ j=1,2 \ldots \quad (5)$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 21, 1989; final revision, Mar. 16, 1990.

Paper No. 91-APM-5.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

and the boundary conditions

$$B_i\phi_j = 0,\ j = 1, 2 \ldots . \tag{6}$$

The solution of equations (5)–(6) consists of a denumerally infinite set of eigenvalues ω_j^2 and the associated eigenfunctions ϕ_j, $j = 1, 2 \ldots$. Since L is self-adjoint, all the eigenvalues are real and positive, and they will be ordered so that $0 < \omega_1 \leq \omega_2 \leq \ldots$. The eigenfunctions can be normalized so as to satisfy

$$\int_0^l m\phi_i\phi_j dx = \delta_{ij},\ i, j = 1, 2 \ldots \tag{7}$$

where δ_{ij} is the Kronecker delta. It follows from equations (5) and (7) that

$$\int_0^l \phi_i L\phi_j dx = \begin{cases} \omega_i^2 & , i = j \\ 0 & , i \neq j \end{cases} \quad i, j = 1, 2 \ldots . \tag{8}$$

Using the eigenfunction expansion theorem, the displacement z can be represented by an infinite sum of the products of modes $\phi_j(x)$ and the corresponding modal amplitudes $a_j(t)$, $j = 1,2, \ldots$. Equations (1) and (4) becomes

$$\ddot{\mathbf{a}} + \omega^2\mathbf{a} = \mathbf{f} + \mathbf{B}\,\mathbf{u},\ t > t_0 \tag{9}$$

$$\mathbf{a}(t_0) = \mathbf{a}_0,\ \dot{\mathbf{a}}(t_0) = \dot{\mathbf{a}}_0 \tag{10}$$

where $\mathbf{a}$, $\mathbf{f}$, $\mathbf{a}_0$, $\dot{\mathbf{a}}_0$, are infinite vectors with their elements given, respectively, by

$$(a_j, a_{oj}, \dot{a}_{oj}) = \int_0^l m(x)[z(x, t), z_o(x), \dot{z}_o(x)]\phi_j(x)dx \tag{11}$$

$$f_j = \int_0^l f\phi_j(x)dx \tag{12}$$

and

$$\Omega^2 = \text{diag}[\omega_j^2],\ \mathbf{B} = [\phi_j(x_i)],\ i = 1, 2 \ldots n,\ j = 1, 2 \ldots . \tag{13}$$

Solution of equations (9)–(10), with the aid of the convolution integral, can be represented by

$$\mathbf{a} = \mathbf{C}(t - t_o)\mathbf{a}_o + \mathbf{S}(t - t_o)\dot{\mathbf{a}}_o + \int_0^t \mathbf{S}(t - \tau)[\mathbf{f}(\tau) + \mathbf{B}\mathbf{u}(\tau)]d\tau \tag{14}$$

where $\mathbf{C}$ and $\mathbf{S}$ are infinite dimensional diagonal matrices of the modal harmonics given by

$$\mathbf{S}(t) = \text{diag}[\omega_j^{-1}\sin\omega_j t],\ \mathbf{C}(t) = \dot{\mathbf{S}}(t) = \text{diag}[\cos\omega_j t],\ j = 1, 2 \ldots . \tag{15}$$

Control Algorithm

A positive performance index J is defined for the current value of time and consists of the kinetic energy, potential energy, and the control effort

$$J(t) = Q_1\int_0^l \frac{1}{2} m\dot{z}^2 dx + Q_2\int_0^l \frac{1}{2} EI(x)\left(\frac{\partial^2 z}{\partial^2 x}\right)^2 dx + \frac{1}{2}\sum_{i=1}^n R_i u_i^2. \tag{16}$$

The functions $Q_1(t)$, $Q_2(t)$ are positive semi-definite and $R_i(t)$ are positive definite. The control forces u_i are determined such that $J(t)$ is a minimum subjected to the constraint of the equations of motion, i.e., equations (1) and (4). This minimization implies that the system will acquire a state as near its state of rest $z = \dot{z} = 0$ as it is possible with limited control effort. Then the variation of $J(t)$ can be given by

$$\delta J = \int_0^l \left(Q_1 m\dot{z}\,\delta\dot{z} + Q_2 EI\frac{\partial^4 z}{\partial x^4}\,\delta z\right)dx + \sum_{i=1}^n R_i u_i \delta u_i \tag{17}$$

where the variations δz and $\delta\dot{z}$ can be written in the form

$$\delta z = \sum_{j=1}^{\infty} \phi_j \delta a_j,\quad \delta\dot{z} = \sum_{j=1}^{\infty} \phi_j \delta\dot{a}_j. \tag{18}$$

These, together with the relation given by equations (7)–(8), yield the modal equivalent of equation (17) i.e.,

$$\delta J = Q_1\dot{\mathbf{a}}^T\,\delta\dot{\mathbf{a}} + Q_2\mathbf{a}^T\,\Omega^2\delta\mathbf{a} + \mathbf{u}^T\,\mathbf{R}\,\delta\mathbf{u} \tag{19}$$

where

$$\mathbf{R} = \text{diag}[R_i],\ i = 1, \ldots . n. \tag{20}$$

From equation (14), the variation of $\delta\mathbf{a}$ and $\delta\dot{\mathbf{a}}$ is

$$\delta\mathbf{a} = \int_{t_0} \mathbf{S}(t - \tau)\mathbf{B}\,\delta\mathbf{u}(\tau)d\tau \tag{21}$$

$$\delta\dot{\mathbf{a}} = \int_{t_0} \mathbf{C}(t - \tau)\mathbf{B}\,\delta\mathbf{u}(\tau)d\tau. \tag{22}$$

Substitution of equations (21)–(22) into (19) yields

$$\delta J = Q_1\dot{\mathbf{a}}^T\int_{t_0}^t \mathbf{C}(t - \tau)\mathbf{B}\,\delta\mathbf{u}(\tau)d\tau + Q_2\mathbf{a}^T\Omega^2\int_{t_0}^t \mathbf{S}(t - \tau)\mathbf{B}\,\delta\mathbf{u}(\tau)d\tau + \mathbf{u}^T\mathbf{R}\,\delta\mathbf{u}. \tag{23}$$

The control force in the time domain $t - t_o$ can then be determined by setting δJ to zero for arbitrary δu, and is given by (Tadjbakhsh and Su, 1989)

$$\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^T\{Q_1\mathbf{S}(t - t_o)\dot{\mathbf{a}} + Q_2[\mathbf{I} - \mathbf{C}(t - t_o)]\mathbf{a}\},\ t > t_o. \tag{24}$$

In practical situations, only a few of the lowest modes designated by the indices 1, 2 . . N_c can be controlled. We denote them by

$$\mathbf{a}_c = [a_1, a_2, a_3, \ldots . a_{N_c}]^T. \tag{25}$$

The remaining modes will be denoted by $\mathbf{a}_r$. In a corresponding manner, vectors and matrices will be partitioned as

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_c \\ \mathbf{a}_r \end{bmatrix},\quad \mathbf{f} = \begin{bmatrix} \mathbf{f}_c \\ \mathbf{f}_r \end{bmatrix},\quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_c \\ \mathbf{B}_r \end{bmatrix},\quad \mathbf{B} = \begin{bmatrix} \mathbf{S}_c & 0 \\ 0 & \mathbf{S}_r \end{bmatrix} \tag{26}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_c & 0 \\ 0 & \mathbf{C}_r \end{bmatrix}\ \mathbf{I} = \begin{bmatrix} \mathbf{I}_c & 0 \\ 0 & \mathbf{I}_r \end{bmatrix},\quad \Omega^2 = \begin{bmatrix} \Omega_c^2 & 0 \\ 0 & \Omega_r^2 \end{bmatrix}.$$

Then equation (24) determines a truncated control force $\mathbf{u}_c$

$$\mathbf{u}_c = -Q_1\mathbf{D}\,\dot{\mathbf{a}}_c - Q_2\mathbf{E}\,\mathbf{a}_c \tag{27}$$

where

$$\mathbf{D} = \mathbf{R}^{-1}\mathbf{B}_c^T\mathbf{S}_c(t - t_0),\ \mathbf{E} = \mathbf{R}^{-1}\mathbf{B}_c^T[\mathbf{I}_c - \mathbf{C}_c(t - t_0)]. \tag{28}$$

Meirovitch and Baruh (1985) developed spatial modal filters for reconstruction of the modal states $\mathbf{a}_c$ and $\dot{\mathbf{a}}_c$. Using orthogonality of eigenfunctions they showed that the problem of observer spillover that was discussed by Balas (1978) can be circumvented. Assuming that the effect of time delay due to computation on the system is small enough to be negligible, the behavior of the controlled modes are determined form

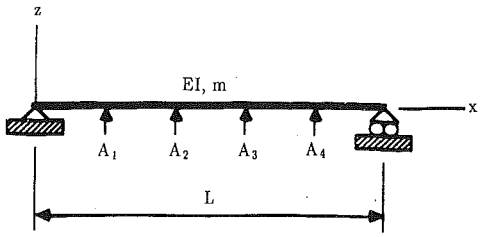
$$\ddot{\mathbf{a}}_c + Q_1\mathbf{B}_c\mathbf{D}\,\dot{\mathbf{a}}_c + (\Omega_c^2 + Q_2\mathbf{B}_c\,\mathbf{E})\mathbf{a}_c = \mathbf{f}_c. \tag{29}$$

The residual modes are subject to control spillover and are determined from

$$\ddot{\mathbf{a}}_r + \Omega_r^2\mathbf{a}_r = \mathbf{f}_r - Q_1\mathbf{B}_r\mathbf{D}\,\mathbf{a}_c - Q_2\mathbf{B}_r\mathbf{E}\,\mathbf{a}_c. \tag{30}$$

The elements of gain matrices consist of modal harmonics $\cos\omega_j(t - t_o)$ and $\sin\omega_j(t - t_o)$ which initially, i.e., near $t = t_o$, are non-negative. As t increases, the higher harmonics begin to change sign thus creating negative damping and reducing the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



System Parameters:

L : 60.m

EI : $5 * 10^5$ N/m^2

m : 1. kg/m

x_1 = 13. m

x_2 = 25. m

Control Parameters:

Q_1 : 5000.

Q_2 : 5000.

R_i : 1., i=1,2,3,4

x_3 = 35. m

x_4 = 47. m

Fig. 1 Simply-supported beam with four discrete actuators, A_i, $i = 1, \ldots 4$

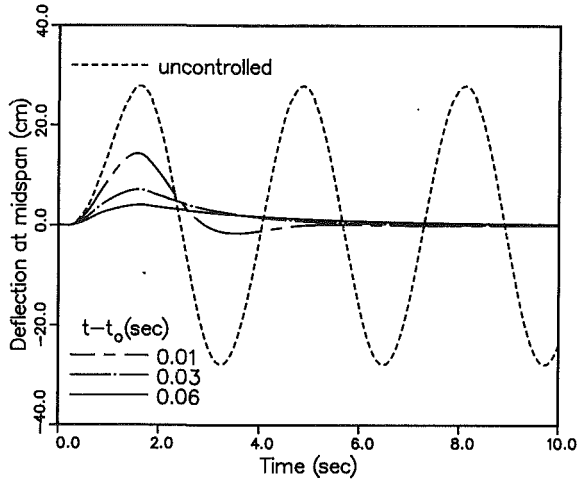


Fig. 3 The midspan deflection of the beam subjected to a moving load traveling at 37 m/s speed, $t - t_o = 0.01$, 0.03, and 0.06 seconds

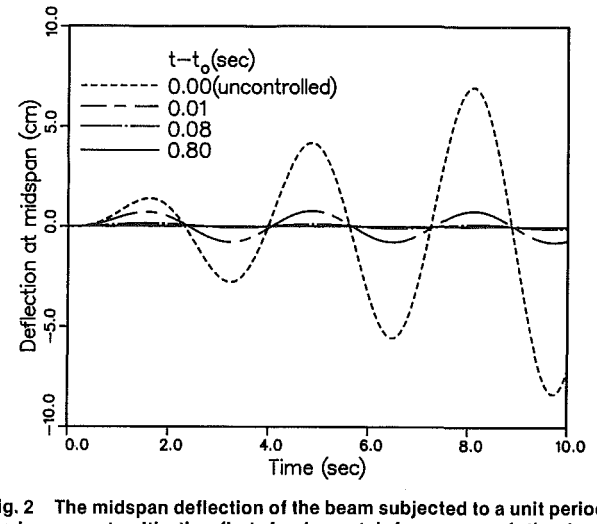


Fig. 2 The midspan deflection of the beam subjected to a unit periodic load resonant with the first fundamental frequency of the beam, $t - t_o = 0.01$, 0.08, 0.80 seconds

stiffness of system. It thus becomes clear that the optimal algorithm derived here is valid for an interval of time during which occurrence of negative damping or reduction of stiffness have not caused amplification of the response of those higher modes that participate significantly in the dynamic of the system. Therefore, the control interval, i.e., $t - t_o$ should not exceed one half of the period of the highest controlled mode, i.e.,

$$(t - t_o) \leq \frac{\pi}{\omega_{N_c}}. \tag{31}$$

Control of a Simply-Supported Beam

Consider a uniform beam hinged at both ends and controlled by four actuators, Fig. 1. The corresponding location of the actuators, x_i, $i = 1, 2, 3, 4$, are 13, 25, 35, and 47 m from the left end of the beam, respectively. We assume that the length of the beam is 60 m, its bending stiffness EI is 5×10^5 N-m^2 and its mass per unit length, m, is 1.0 Kg(m). Then the eigenvalues ω_i^2 and normalized eigenfunctions are given by

$$\omega_i^2 = 0.039(i\pi)^4, \quad \phi_i(x) = 0.183 \sin(i\pi/60), \; i = 1, 2 \ldots .$$

Table 1

Mode	Max.$(t - t_o)$, sec	Mode	Max.$(t - t_o)$, sec
1	1.621	6	0.045
2	0.405	7	0.033
3	0.180	8	0.025
4	0.101	9	0.020
5	0.065	10	0.016

The control parameters are assigned fixed values $Q_1 = Q_2 = 5000$ and $R_i = 1$, $i = 1, 2, 3, 4$. The lowest ten modes of the beam constitute the controlled and the uncontrolled modes.

In the first example, the beam is subjected to a unit cyclic load resonant with the first mode of the beam acting at the midspan. The external excitation in equation (1) is expressed as $f(x, t) = \sin(1.95t) \cdot \delta(x - 30)$. The control intervals are 0.01, 0.08, and 0.80 seconds and only the first mode is subjected to control. Control of additional modes is not called for as they are not being excited. The results that are presented refer to the deflection at midspan of the beam as shown in Fig. 2. It can be seen that, unlike the uncontrolled beam, the response remains bounded. From Table 1 the maximum control gain for the first mode is approximately 0.81 seconds. Thus, the beam is almost completely controlled, without any oscillations, by setting the control interval, i.e., $t - t_0$, equal to 0.80 seconds.

In the second example the beam is subjected to a 20N load that travels along the beam with a speed of 37 m/sec. Now the lowest five modes are controlled and the next five modes are uncontrolled ($N_c = N_r = 5$). The results corresponding to control intervals of 0.01, 0.03, and 0.06 seconds are shown in Fig. 3. It is clear that the controllers suppress the vibration within 10 seconds.

The above two examples show that, within the restriction imposed by equation (29), larger control intervals bring about more effective controls. From a practical point of view a digital control system of the controller is desirable. This can be achieved by a discretized version of the continuous time system expressed in equations (27) and (28), cascaded with Meirovitch's spatial modal filters and a zero-order hold analog to digital converter (Öz, *H*, et al., 1980).

Hence, it is conjectured that for given N_c and N_r, an optimal control interval exits that takes into account the two issues of stability and effectiveness of the control. The optimal size of the control interval is postponed to future investigations.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

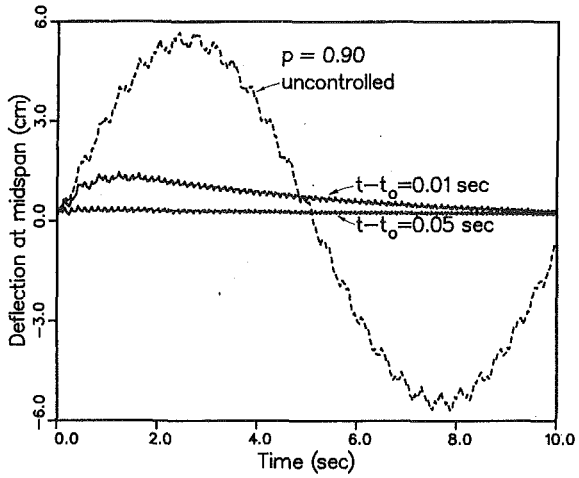


Fig. 4 The midspan deflection of the controlled beam column subjected to a unit impulse at midspan, $p = 0.9$, $t - t_o = 0.01$, 0.05 seconds

Control of Simply-Supported Beam Columns

Beam columns are members that are subjected to both bending and axial compression. They will be unstable when the axial compression reaches certain critical values. For a given beam column with certain specified boundary conditions, theoretically, an infinite number of critical loads and associated buckling mode shapes exist. In this section, we attempt to follow the same procedure derived previously to control the beam columns at and above their buckling load.

Let P denote the axial compressive load applied to a beam in Fig. 1, the eigenvalue problem is given by

$$\frac{\partial^2}{\partial x^2}\left(EI(x)\,\frac{\partial^2\phi_j}{\partial x^2}\right)+P\,\frac{\partial^2\phi_j}{\partial x^2}=m\omega_j^2\phi_j,\ j=1,\ 2,\ 3\ \ldots \tag{32}$$

with the boundary conditions

$$\phi_j=\frac{\partial^2\phi_j}{\partial x^2}=0 \text{ for } x=0,\ l \tag{33}$$

that hold for the simply-supported case.

The normalized eigenfunctions for equations (32)–(33) for a prismatic beam can be written in the form

$$\phi_j=\sqrt{(2/ml)}\sin\frac{j\pi x}{l},\ j=1,\ 2,\ \ldots. \tag{34}$$

As a result of substitution of equation (34) into equation (32), we have

$$\omega_j=\pm\begin{cases}j\mu\gamma_j & \text{if } p<j,\ j=1,\ 2,\ \ldots.\\ ij\mu\gamma_j & \text{if } p>j,\ j=1,\ 2,\ \ldots.\end{cases} \tag{35}$$

where

$$p=\frac{P}{P_E},\quad \gamma_j=\sqrt{j^2-p} \tag{36}$$

$$\mu=\left(\frac{\pi^2}{l^2}\right)\sqrt{\frac{EI}{m}},\quad i=\sqrt{-1}$$

and P_E denotes the Euler load of the beam column, i.e., $P_E=\pi^2EI/l^2$.

When p is smaller than 1, i.e., the axial load P is smaller than P_E, the beam is stable and the control law expressed in equation (24) is still valid with the ω_j, $j=1,\ 2,\ 3\ \ldots$ given by equation (35). When the beam is in critical condition, i.e., $p=1$, then ω_1 is zero. However, because $\lim_{p\to 1}\sin(\omega_1 t)/\omega_1=t$, and $\lim_{p\to 1}\cos(\omega_1 t)=1$, one obtains from equation (15)

$$\begin{aligned}\mathbf{S}(t)&=\text{diag}[t,\ \omega_j^{-1}\sin(\omega_j t)],\quad j=2,\ 3\\ \mathbf{C}(t)&=\text{diag}[1,\ \cos(\omega_j t)],\quad j=2,\ 3,\ \ldots.\end{aligned} \tag{37}$$

Following the same procedure described in previous sections, one can obtain the same control law that is already shown in equations (27)–(28), in which $\mathbf{S}$ and $\mathbf{C}$ are now defined by equation (37).

Similarly, when $1<p<2$, then the control force and the system behavior can be given by equations (27)–(30) wherein $\mathbf{S}$, $\mathbf{C}$ can be written in the form

$$\begin{aligned}\mathbf{S}(t)&=\text{diag}[\,|\omega_1|^{-1}\sinh|\omega_1|t,\ \omega_j^{-1}\sin(\omega_j t)],\ j=2,\ 3,\ \ldots.\\ \mathbf{C}(t)&=\text{diag}[\cosh|\omega_1|t,\ \cos(\omega_j t)],\ j=2,\ 3,\ \ldots.\end{aligned} \tag{38}$$

For numerical evaluation, the beam with the same parameters as shown in Fig. 1 is used. The system is subjected to an impulse at midspan. The control intervals are set to 0.01 and 0.05 seconds to compare results. The behavior of the beam column is assumed to be adequately represented by its first ten modes and of these the lowest five are controlled. Figure 4 shows the response of the system over a 10-s interval. The axial compressive load ratio $p=P/P_E$ is set to 0.9. The deflection of the beam remains bounded and is accompanied by control force spilling over into the higher modes. The time histories of the commanded forces for the actuator A_2—one of the middle two actuators—are shown in Fig. 5. The larger control interval causes higher beginning values for the force and lower tail end values. Thus it may be concluded that the control is effective at the axial load of 0.9 of buckling load.

Next, we consider the control of a beam column at the buckling case, i.e., $p=1$. In this case, the $\mathbf{E}$ defined in equation (28) can be partitioned as

$$\mathbf{E}=\begin{bmatrix}E_{11}\mathbf{E}_{12}\\ \mathbf{E}_{21}\mathbf{E}_{22}\end{bmatrix}=\begin{bmatrix}\begin{Bmatrix}0\\ \vdots\\ 0\end{Bmatrix}\begin{bmatrix}(1-C_2)\dfrac{B_{21}}{R_1} & \ldots. & (1-C_{N_c})\dfrac{B_{N_c1}}{R_1}\end{bmatrix}\\ \begin{Bmatrix}0\\ \vdots\\ 0\end{Bmatrix}\begin{bmatrix}(1-C_2)\dfrac{B_{2n}}{R_n} & \ldots. & (1-C_{N_c})\dfrac{B_{N_c2}}{R_n}\\ \vdots & & \vdots\\ (1-C_2)\dfrac{B_{2n}}{R_n} & \ldots. & (1-C_{N_c})\dfrac{B_{N_cn}}{R_n}\end{bmatrix}\end{bmatrix} \tag{39}$$

where

$$C_i\colon \cos(\omega_i(t-t_o)),\ i=1\ \ldots.\ N_c.$$

Also, the $\mathbf{B}_c$, $\mathbf{D}$ matrices and the $\mathbf{a}_c$ vector can be partitioned as

$$\mathbf{B}_c=\begin{bmatrix}B_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{bmatrix}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

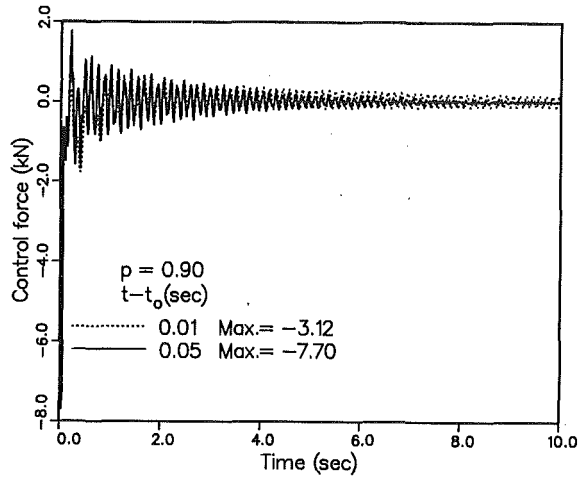


Fig. 5 Commanded force of A_2 actuator for the control of the beam column, $p = 0.9$, $t - t_o = 0.01$, 0.05 seconds

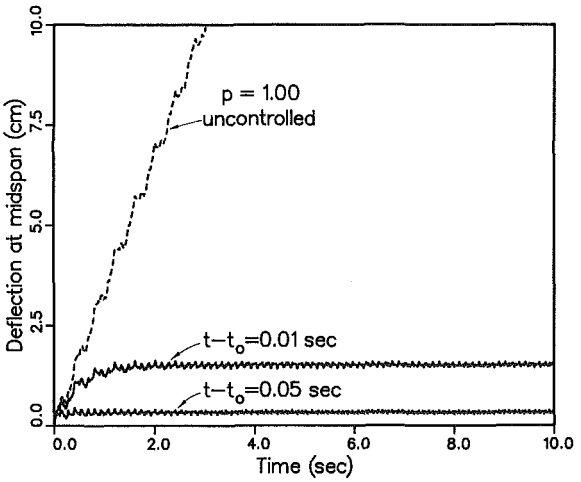


Fig. 6 The midspan deflection of the controlled beam-column subjected to a unit impulse at midspan, $p = 1.0$, $t - t_o = 0.01$, 0.05 seconds

$$\mathbf{D} = \begin{bmatrix} D_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \tag{40}$$

$$\mathbf{a}_c = \begin{Bmatrix} a_1 \\ \mathbf{a}_2 \end{Bmatrix}.$$

Then the first modal equation that corresponds to the buckling mode of the beam column in equation (29) becomes

$$\ddot{a}_1 + Q_1 B_{11} D_{11} \dot{a}_1 = f_1 - Q_1 \mathbf{B}_{12} \mathbf{D}_{12} \dot{\mathbf{a}}_2 - Q_2 (B_{11} \mathbf{E}_{12} + \mathbf{B}_{12} \mathbf{E}_{22}) \mathbf{a}_2 \tag{41}$$

with appropriate initial conditions at t_0. From equation (41), it is clear that the feedback control law can only provide a certain amount of damping to the system and the coupled modal control force acts as a counter force to balance f_1. Because a_1 is coupled with $\mathbf{a}_2$, it is difficult to find the analytical solution for a_1. Therefore, the same numerical procedure employed previously is used to find approximate results. Figure 6 shows the results of the control of the beam column at its buckling load when the control intervals are set to 0.01 and 0.05 seconds. The control spillover is present, and the system acquires a permanent deflection that depends on the control interval in both cases. Different control efforts are pumped into the system for the different control intervals, and this is responsible for the different levels of permanent deformations shown in Fig. 6. As observed from Fig. 7, the commanded force for smaller control intervals have higher initial and lower final values.

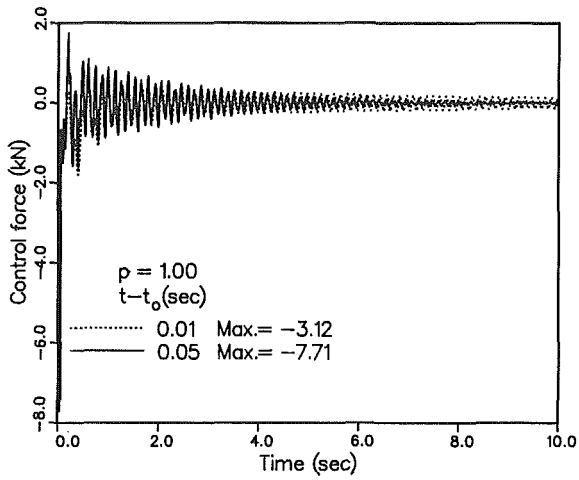


Fig. 7 Commanded force of A_2 actuator for the control of the beam column, $p = 1.0$, $t - t_0 = 0.01$, 0.05 seconds

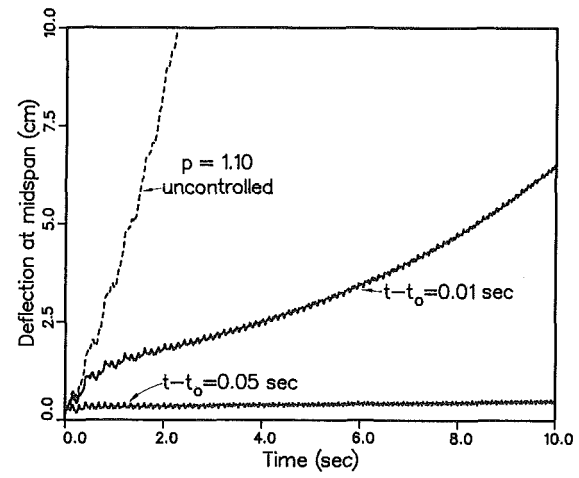


Fig. 8 The midspan deflection of the controlled beam column subjected to a unit impulse at midspan, $p = 1.1$, $t - t_o = 0.01$, 0.05 seconds

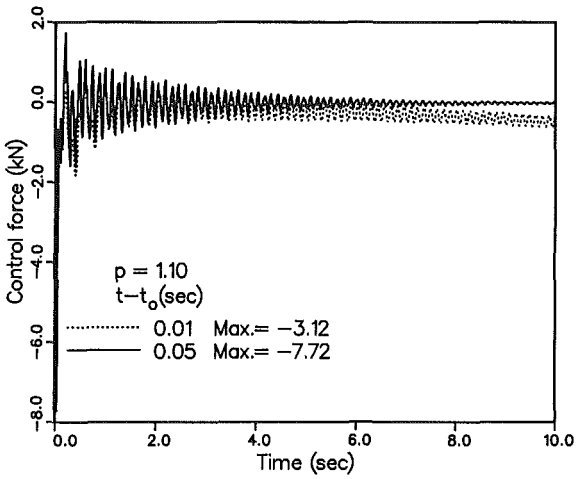


Fig. 9 Commanded force of A_2 actuator for the control of the beam column, $p = 1.1$, $t - t_o = 0.01$, 0.05 seconds

Figures 8 and 9 show the control of the beam column when $p = 1.1$. In Fig. 8, the system appears unstable with deflection

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

growing continually for both control intervals of 0.01 and 0.05 seconds. The deflection of the control interval of 0.05 seconds have a smaller rate of growth than that of the control interval of 0.01 seconds.

Figure 9 shows the A_2 the actuator's commanded force histories for both cases. The importance of the beginning stages of control effort, being a function of the control interval, shows up prominently in these cases. In the case of the short control, i.e., $t-t_0=0.01$ seconds, the commanded force of the actuator A_2 increases with the unbounded response of the controlled system. However, in the long control interval case, i.e., $t-t_0=0.05$ seconds, the commanded force grows with a more moderate rate compared to that of the short interval. Therefore, it is clear that the longer control intervals are better in that they provide slower rate of growth of deflections.

Concluding Remark

The couple-modal control algorithm can successfully be used to suppress the vibration of the beam subjected to a resonant cyclic load and a moving load as shown in this paper. The algorithm is causal, i.e., it only depends on the past history of the system and such that the control interval dominates the system's behavior. The maximum control interval as given by equation (31) is a function of the period of the highest controlled mode. The response of controlled system is significantly affected by the control interval. Therefore, a proper choice of the control interval is necessary for the algorithm. The algorithm can be applied to control of a beam column before buckling. However, the controller can only provide a damping resistance to the buckled mode, and the algorithm is not well suited for controlling a beam after buckling.

Acknowledgment

The work reported here was conducted in the course of research supported by the Army Research Office under Contract DAAL 03-89-K-0002 with Rensselaer Polytechnic Institute. This support is gratefully acknowledged.

References

Balas, M. J., 1978, "Active Control of Flexible System," Journal of Optimization Theory and Applications, Vol. 25, pp. 415–436.

Kirk, D., 1970, *Optimal Control Theory, An Introduction*, Prentice-Hall, Englewood Cliffs, New Jersey.

Luenberger, D., 1968, "An Introduction to Observers, *IEEE Trans. Automat. Control.*, Vol. AC-16, No. 6, p. 316–353.

Meirovitch, L., and Öz, H., 1980, "Modal-Space Control of Distributed Gyroscopic Systems," *Journal of Guidance and Control*, Vol. 3, No. 2, pp. 140–150.

Meirovitch, L., and Baruh, H., 1982, "Control of Self-Adjoint Distributed Parameter Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 1, pp. 60–66.

Meirovitch, L., and Baruh, H., 1983, "Robustness of the Independent Modal-Space Control Method," *Journal of Guidance, Control, and Dynamics*, Vol. 6, No. 1, pp. 20–25.

Meirovitch, L., and Baruh, H., 1985, "The Implementation of Modal Filters for Control of Structures," *Journal of Guidance, Control, and Dynamics*, Vol. 8, No. 6, pp. 707–716.

Meirovitch, L., 1987, "Some Problems Associated with Control of Distributed Structures," *Journal of Optimization Theory and Applications*, Vol. 54, No. 1, pp. 1–21.

Öz, H., Meirovitch, L., and Johnson, Jr., C. R., 1980, "Some Problems Associated with Digital Control of Dynamic Systems," *Journal of Guidance and Control*, Vol. 3, No. 6, pp. 523–528.

Öz, H., and Meirovitch, L., 1980, "Optimal Modal-Space Control of Flexible Gyroscopic Systems," *Journal of Guidance and Control*, Vol. 3, No. 3, pp. 218–226.

Simon, J., and Mitter, S. A., 1968, "A Theory of Modal Control," *Information and Control*, Vol. 13, pp. 316–353.

Tadjbakhsh, I. G., and Su, Y. A., 1989, "Optimal, Coupled-Modal Control of Distributed Parameter Systems with Discrete Actuators," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, No. 4, pp. 941–946.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Accurate Analytical-Type Solutions for the Free Vibration of Simply-Supported Parallelogram Plates

D. J. Gorman
Professor,
Department of Mechanical Engineering,
University of Ottawa,
Ottawa, Ontario K1N 6N5, Canada

A comprehensive study of the free vibration of simply-supported parallelogram plates is conducted. Solutions are obtained by utilizing the superposition method and by taking advantage of symmetry inherent in the problem. Toward this end a new alternating Lévy-type solution is introduced. Verification tests are conducted by comparing computed eigenvalues with those of rhombic plates in the special case where all plate edges are of equal length. Eigenvalues are stored for eight vibration modes and for a wide range of plate geometry.

1 Introduction

A search of the literature quickly reveals that little accurate information has been published up to this time with regard to the free vibration frequencies and mode shapes of parallelogram plates. There is no doubt that failure of researchers to fully address this problem takes it roots in recognition of the extremely formidable task of trying to satisfy the free vibration governing differential equation throughout the plate while at the same time satisfying prescribed boundary conditions along the plate edges. One could easily be tempted to conclude that achievement of accurate solutions for problems of this type is just not mathematically possible.

In this paper a new approach is taken wherein the superposition method, as developed by the author in numerous earlier publications, is brought to bear on the problem of analyzing general parallelogram plates with simple support conditions imposed along all edges. The problem is greatly simplified by taking advantage of symmetry and antisymmetry conditions which characterize all possible free vibration mode shapes of the problem at hand. Satisfaction of boundary conditions is achieved by introducing a new type of alternating series for Lévy-type solutions. Finally, before presenting a rather wide range of eigenvalues for various plate geometries, it is shown how accurate solutions for the much simpler problem of free vibration of rhombic plates as published by the author in an earlier paper, are utilized to verify the correctness of the solution for the parallelogram plate problem.

2 Formulation of the Problem

A study of the simply-supported parallelogram plate of Fig. 1 permits the viewer to arrive at certain conclusions concerning the possible free vibration mode shapes for this plate. It will be evident that all modes will possess a degree of symmetry or antisymmetry with respect to the central axis of the plate as shown in the figure.

For all modes possessing a degree of symmetry with respect to this axis (including the fundamental mode), we may write for the plate lateral displacement

$$W(\xi,\eta) = W(-\xi,-\eta). \tag{1}$$

All other modes will possess a degree of antisymmetry with respect to the axis and for these modes we may write,

$$W(\xi,\eta) = -W(-\xi,-\eta). \tag{2}$$

We shall refer to all modes henceforth as being of the symmetric or antisymmetric type.

The problem is greatly simplified, in view of the above observations, as we shall now find that we need analyze only half of the parallelogram plate.

Consider the parallelogram plate of Fig. 2 undergoing symmetric mode vibration, let us say first mode vibration. We focus attention on the central rectangular portion of the plate as shown in heavy outline. The bending moments acting on each edge of this rectangular portion, as a result of interfacing with the triangular portions of the remainder of the parallelogram, will have a distribution somewhat as shown in the figure. It will be apparent that if these moments were expanded in a Fourier sine series the first Fourier components of each moment would have the same sign, i.e., if one was positive the other would also be positive. The second components, however, would have opposite signs. In fact, these signs would

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 25, 1989; final revision, Mar. 16, 1990.

Paper No. 91-APM-6.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

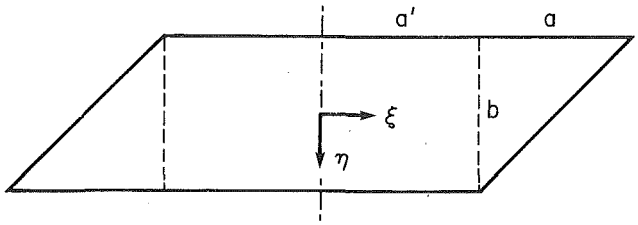


Fig. 1 General view of the simply-supported parallelogram plate

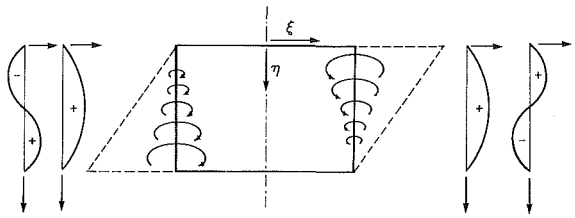


Fig. 2 Parallelogram plate in symmetric-type mode vibration. Fourier components of bending moments at rectangle-triangle interfaces are shown.

alternate as we move on up to higher numbers of terms in the series representing these harmonic bending moments. The first two Fourier components for each distributed moment are shown schematically in the figure.

It will be evident that associated with first, third, and fifth Fourier components, etc., in the moment distributions, slip shear conditions would prevail along the central η-axis of the plate. That is, slope of the plate taken normal to the direction of the η-axis would be everywhere zero along this axis and, furthermore, plate transverse shear taken along this axis would be everywhere zero. Associated with the even-numbered Fourier components we would have a condition of zero lateral displacement along the η-axis and, in addition, bending moments acting along the axis would be everywhere zero. These observations will prove to be very important in developments to follow shortly.

In Fig. 3 a typical distribution of the bending moments acting along the extremities of the rectangular portion of the plate is shown for the plate in antisymmetric type mode vibration. Here it is seen that the Fourier components of the moments at each edge of the rectangular portion of the plate change sign in a manner exactly opposite to that described above for symmetric mode vibration. Accordingly, for the first, third, and fifth Fourier components, etc., a condition of zero lateral displacement and zero bending moment will prevail along the η-axis. Conversely, slip shear conditions will prevail along the same axis when the even numbered Fourier components are applied.

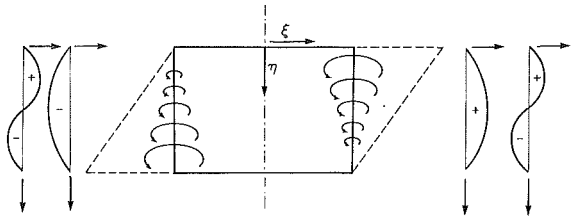


Fig. 3 Parallelogram plate in antisymmetric-type mode vibration. Fourier components of bending moments at rectangle-triangle interfaces are shown.

In the theoretical model considered here, only the right-hand segment of the rectangular portion of the parallelogram plate (Figs. 2 and 3) as well as the adjoining right triangular segment are included in the analysis. Each segment is analyzed separately and conditions of continuity are enforced along the common boundary. Appropriate boundary conditions are enforced along the edge of the rectangular segment corresponding to the η-axis of Figs. 2 and 3 so that conditions dictated by symmetry, or antisymmetry, are satisfied. It should be pointed out that the same conclusions arrived at through considering bending moment distributions in Figs. 2 and 3 could have been arrived at by considering transverse shear forces instead.

Mathematical Procedure. We begin by analyzing the right triangular segment of the plate as discussed above. The solution for this segment is achieved by the superposition method as described in a number of earlier publications. The publication most relevant to the triangular plate segment under study here is by Gorman (1989) which deals with the analysis of right triangular plates with one free edge. The reader may wish to examine this reference as many of the details presented therein will not be repeated here.

Analysis of the right triangular segment is achieved by superimposing solutions for the six rectangular building blocks of Fig. 4. All nondriven edges of these building blocks have simple edge support as indicated by a straight heavy line, or slip shear edge conditions (as discussed earlier) which are indicated by a straight line with two adjacent small circles. The first and third building blocks have edges which are driven by a distributed harmonic bending moment or rotation, with lateral displacement forbidden along these driven edges. The fifth building block is driven in a similar manner except that, instead of forbidding displacement, vertical edge reaction is forbidden along the driven edge. All other driven edges are driven by a distributed harmonic vertical edge reaction with bending moment forbidden along the edge.

A solution is provided here for the first building block only. The solution is taken in the form proposed by Lévy as

Nomenclature

a, b = edge lengths of right triangular portions of parallelogram plate
a' = edge length of half of rectangular portion of parallelogram plate
D = flexural rigidity of plate = $Eh^3/12(1 - \nu^2)$
E = Young's modulus of plate material
W = plate lateral deflection divided by a
ν = Poisson's ratio of plate material
ρ = mass of plate per unit area
ω = circular frequency of plate vibration
ϕ = aspect ratio of right triangular portions of plate = b/a
ϕ' = aspect ratio of half of rectangular portion of plate = b/a'
ϕ_r = aspect ratio of rhombic plate
ξ = distance along right triangular portion of plate, or rectangular portion of half plate, divided by a, or a', respectively
η = distance along common boundary of right triangular and rectangular portions of plate divided by b
λ^2 = plate eigenvalue = $\omega a^2 \sqrt{\rho/D}$
λ^{*2} = alternate formulation of eigenvalue = $\omega b^2 \sqrt{\rho/D}$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

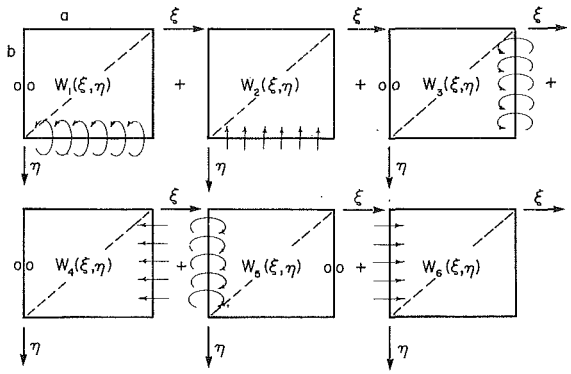


Fig. 4 Rectangular building blocks used to generate triangular portion of parallelogram

$$W_1(\xi,\eta)=\sum_{m=1,3,5}^{\infty} Y_m(\eta)\cos\frac{m\pi\xi}{2} \tag{3}$$

where the time-varying component is not included.

We consider the building block to be driven by a distributed harmonic rotation whose amplitude is expanded in the same Fourier series utilized in equation (3) as follows:

$$\frac{\partial W_1(\xi,\eta)}{\partial\eta}\Big|_{\eta=1}=\sum_{m=1,3,5}^{\infty} E_m\cos\frac{m\pi\xi}{2} \tag{4}$$

where the E_m's are coefficients to be determined.

Solutions for the functions $Y_m(\eta)$ of equation (3) are well known and contain four constants to be determined. Two of the constants are immediately deleted in view of the boundary conditions imposed along the edge, $\eta = 0$. The other constants are determined by satisfying conditions imposed along the driven edge. It is readily shown that we may write (Gorman, 1989)

$$Y_m(\eta)=E_m\left\{\theta_{11m}\frac{\sinh\beta_m\eta}{\sinh\beta_m}+\theta_{13m}\frac{\sin\gamma_m\eta}{\sin\gamma_m}\right\} \tag{5}$$

for $\lambda^2 > \left(\frac{m\pi}{2}\right)^2$, and

$$Y_m(\eta)=E_m\left\{\theta_{22m}\frac{\sinh\beta_m\eta}{\sinh\beta_m}+\theta_{23m}\frac{\sinh\gamma_m\eta}{\sinh\gamma_m}\right\} \tag{6}$$

for $\lambda^2 < \left(\frac{m\pi}{2}\right)^2$.

Furthermore, it is shown that

$$\theta_{11m}=\frac{-1}{(\beta_m^2+\gamma_m^2)},\quad \theta_{13m}=-\theta_{11m}$$

$$\theta_{22m}=\frac{-1}{(\beta_m^2-\gamma_m^2)},\quad \theta_{23m}=-\theta_{22m}$$

where $\beta_m = \phi\sqrt{\lambda^2+\left(\frac{m\pi}{2}\right)^2}$, and $\gamma_m = \phi\sqrt{\lambda^2-\left(\frac{m\pi}{2}\right)^2}$, or $\phi\sqrt{\left(\frac{m\pi}{2}\right)^2}$, whichever is real.

Solutions for the remaining building blocks of Fig. 4 are easily achieved in a similar manner. Some solutions can be extracted from others through a simple interchange of variables. Before utilizing these building block solutions to construct the eigenvalue matrix, we must develop solutions for the rectangular plate segment as shown in the right-hand side of Fig. 2. Let the coordinates ξ and η be so nondimensionalized that they run between the limits of zero and one for this segment.

The solution for the segment is taken in the form proposed by Lévy as

$$W(\xi,\eta)=\sum_{n=1,2}^{\infty} Y_n(\xi)\sin n\pi\eta. \tag{7}$$

It is recalled that the solution for $Y_n(\xi)$ is written as,

$$Y_n(\xi)=A_n\cosh\beta_n\xi+B_n\sinh\beta_n\xi+C_n\cos\gamma_n\xi+D_n\sin\gamma_n\xi \tag{8}$$

for $\lambda^2 > (n\pi)^2$, and

$$Y_n(\xi)=A_n\cosh\beta_n\xi+B_n\sinh\beta_n\xi+C_n\cosh\gamma_n\xi+D_n\sinh\gamma_n\xi \tag{9}$$

for $\lambda^2 < (n\pi)^2$, where β_n and γ_n are as given above for equations (5) and (6) except that n replaces $m/2$ λ^2 is now replaced by $\lambda^{*2} = \omega b^2\sqrt{\rho/D}$, and the plate aspect ratio equals b/a'.

We first consider symmetric-type vibration modes for the parallelogram plate. In order to satisfy conditions along the edges, $\xi = 0$, for these modes, it will be evident, in view of our earlier discussion, that we must delete the sinh and sine terms from equations (8) and (9) when the subscript n takes on odd integer values. Conversely, we must delete the cosh and cosine terms when n is even. As the solution for equation (7) we may therefore write, for $n = 1, 3, 5 \ldots$ etc.,

$$Y_n(\xi)=A_n\cosh\beta_n\xi+B_n\cos\gamma_n\xi$$

or

$$A_n\cosh\beta_n\xi+B_n\cosh\gamma_n\xi \tag{10}$$

and for $n = 2, 4, 6 \ldots$ etc.,

$$Y_n(\xi)=A_n\sinh\beta_n\xi+B_n\sin\gamma_n\xi$$

or

$$A_n\sinh\beta_n\xi+B_n\sinh\gamma_n\xi. \tag{11}$$

The solution for equation (7), when antisymmetric mode vibration is considered, is easily extracted from equations (10) and (11). For these latter modes the rules are applied in the exact opposite sense, and we obtain, for $n = 1, 3, 5 \ldots$ etc.,

$$Y_n(\xi)=A_n\sinh\beta_n\xi+B_n\sin\gamma_n\xi$$

or

$$A_n\sinh\beta_n\xi+B_n\sinh\gamma_n\xi \tag{12}$$

and for $n = 2, 4, 6 \ldots$ etc.,

$$Y_n(\xi)=A_n\cosh\beta_n\xi+B_n\cos\gamma_n\xi$$

or

$$A_n\cosh\beta_n\xi+B_n\cosh\gamma_n\xi. \tag{13}$$

4 Construction of Eigenvalue Matrix and Generation of Eigenvalues

A schematic representation of the eigenvalue matrix, based on a three-term expansion of the driven edge rotations, etc., is provided in Fig. 5. The upper two segments of the matrix are arrived at in a manner identical to that described in Gorman (1989). Interest is focused on the right triangular portion of the building blocks of Fig. 4, lying above the broken line diagonals. The contributions of each of the first six building blocks in the figure toward displacement, δ, along the diagonal shown in the small rectangle to the right of the figure are expanded in a *cosine* series. We demand that the summation of the contributions toward each Fourier term in the expansion should vanish, thereby arriving at a set of simultaneous homogenous algebraic equations relating the driving coefficients, E_m, E_n, etc. (Gorman, 1989). In the second segment a similar set of equations arise because of the need to satisfy the condition of zero net bending moment along the same diagonal. Following established procedures a set of $2k$ equations is written for each segment where k is the number of terms utilized in the driving force expansions. In this way the total number

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

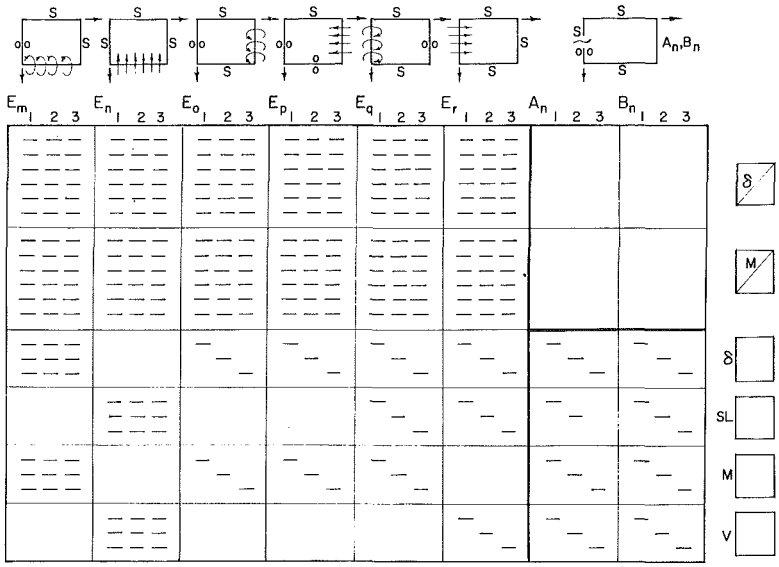


Fig. 5 Schematic representation of eigenvalue matrix based on three term expansions. Simple support or slip-shear conditions may prevail at left edge of final building block depending on subscript "*n*" and type of mode under study.

of equations eventually written will equal the number of unknowns.

Moving down toward the final four segments in the matrix we see that their associated homogeneous equations arise because of the need to satisfy the conditions of continuity at the common boundary between the right triangular plate segment and the rectangular plate segment as discussed earlier. These continuity conditions involve displacement, slope, bending moment, and vertical edge reaction as indicated on the right-hand side of Fig. 5. The contributions of each building block toward displacement, etc., are expanded in a sine series and the coefficients are constrained so as to insure continuity. The elements in the matrix related to the unknown coefficients A_n and B_n will, of course, depend on whether we are studying symmetric or antisymmetric-type vibration modes and whether the subscript n is odd or even. All of the necessary applicable rules have already been discussed.

Eigenvalues are obtained following standard procedures. A trial eigenvalue is selected, the eigenvalue matrix is generated, and the determinant of the matrix is established and stored. The trial eigenvalue is then augmented and the procedure is repeated until the determinant changes sign. The actual eigenvalue, for which the determinant vanishes, is then established. The number of terms in the series, k, may be increased in order to check convergence. After setting one of the nonzero coefficients equal to unity the mode shape of the parallelogram plate is computed.

5 Presentation of Computed Results

An important step which must be taken before computing eigenvalues for the parallelogram plate involves setting up an orderly scheme for storage of these eigenvalues. The scheme proposed here is shown schematically in Fig. 6. The figure outlined represents one half (the right half) of a parallelogram plate. Referring to the rectangular portion of this half plate, the quantity $1/\phi'$ represents the ratio of the width of this portion to its edge length b. Computations have been carried out for values of this ratio varying between the limits 0.4 and 1.6. For the right triangular portion of the plate the quantity $1/\phi$ represents the ratio of the width of the base of the triangle to the edge length b. This ratio is allowed to vary between the limits 0.4 and 1.2. All stored eigenvalues are nondimensionalized with respect to the dimension b, i.e., we write $\lambda^{*2} = \omega b^2 \sqrt{\rho/D}$. All possible combinations of values of the parameters $1/\phi'$ and $1/\phi$ are included in the stored tables and the first four modes for symmetric mode-type vibration and antisymmetric mode-type vibration are computed and stored.

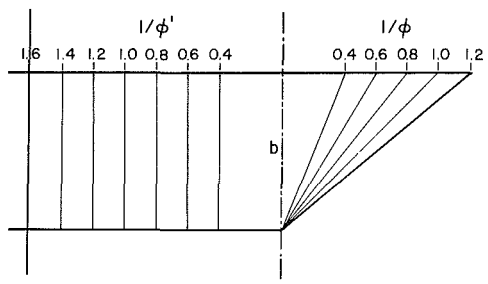


Fig. 6 Geometries of parallelogram plates for which eigenvalues are stored

Before examining the stored results we wish to conduct some tests to verify the validity of the theoretical model. In an earlier paper dealing with the much simpler problem of rhombic plate free vibration, (Gorman, 1988), it was pointed out that since the rhombic plate constituted a special case of the parallelogram plate (a parallelogram plate with all edges of equal length) the results of the rhombic plate study could provide verification points for verifying parallelogram plate studies. Free vibration of rhombic plates was also studied in Kaul and Cadambe (1956). The equivalence of these two types of plates is illustrated in Fig. 7. Advantage of this phenomena has been taken in the work reported here and in Tables 1 and 2 a comparison is made between eigenvalues for rhombic plates, extracted from Gorman (1988), and eigenvalues computed using the much more sophisticated analysis described here for corresponding parallelogram plates. The quantity ϕ_r is the aspect ratio of the rhombic plate as defined in Gorman (1988), but all eigenvalues are nondimensionalized with respect to the length b as defined in Fig. 6. It should be pointed out that all fully symmetric and fully antisymmetric modes of the rhombic plate as defined in Gorman (1988) correspond to symmetric-type modes of the parallelogram plate (Table 1). Conversely, all symmetric-antisymmetric modes of the rhombic plate correspond to antisymmetric-type modes of the parallelogram plate (Table 2).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

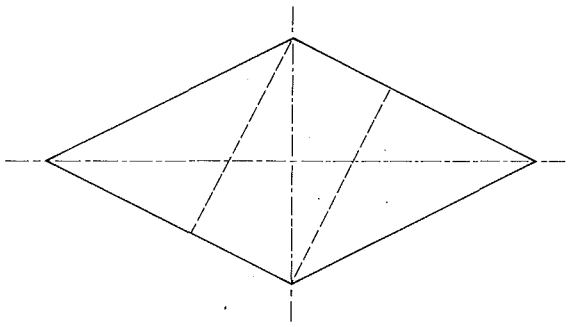


Fig. 7 Demonstration of equivalence of rhombic plates and certain parallelogram plates

Table 1 Comparison between eigenvalues inferred from rhombic plate results and those computed for equivalent parallelogram plate (in parentheses)

Fully symmetric modes (rhombic plate)

Mode:	$\phi_R = 1.5$	$\phi_R = 2.0$
1	3.318 (3.317)	10.15 (10.14)
2	11.90 (11.90)	32.15 (32.13)
3	18.00 (18.00)	58.80 (58.79)

Fully antisymmetric modes (rhombic plate)

Mode:	$\phi_R = 1.5$	$\phi_R = 2.0$
1	16.48 (16.48)	49.97 (49.97)
2	31.53 (31.53)	89.78 (89.79)
3	44.16 (44.16)	134.4 (134.4)

Table 2 Comparison between eigenvalues inferred from rhombic plate results and those computed for equivalent parallelogram plate (in parentheses)

Modes symmetric about ξ-axis, antisymmetric about η-axis (rhombic plate) (Gorman 1988)

Mode:	$\phi_R = 1.5$	$\phi_R = 2.0$
1	7.346 (7.345)	20.63 (20.62)
2	17.78 (17.78)	46.26 (46.25)
3	28.54 (28.54)	79.81 (79.81)

Modes symmetric about η-axis, antisymmetric about ξ axis (rhombic plate) (Gorman 1988)

Mode:	$\phi_R = 1.5$	$\phi_R = 2.0$
1	9.240 (9.240)	30.20 (30.19)
2	22.92 (22.92)	68.11 (68.10)
3	29.97 (29.96)	99.49 (99.48)

It is seen that there is excellent agreement between the two sets of eigenvalues as tabulated in the tables. Twenty terms have been used in representing the building block solutions for these studies. The high level of agreement attests to the accuracy of the methods utilized in both plate studies.

As further verification of the analysis, we utilize the fact that for low values of the aspect ratio ϕ', we should approach cylindrical bending in fundamental mode vibration. A study of the first row of Table 3 indicates this to be the case. It will be noted that the eigenvalue approaches the correct limit of π^2 as the parameter ϕ' increases.

Computed eigenvalues for the general parallelogram plate are stored in Tables 3 and 4 for symmetric and antisymmetric type mode vibration, respectively. All computations have been carried out with 20 terms in the building block solutions and it is believed that in most cases four-digit accuracy has been achieved.

Mode shapes for the first symmetric-type mode, and the first antisymmetric-type mode, of free vibration of a full parallelogram plate are presented in Figs. 8 and 9, respectively. Even though all edges are of equal length the parallelogram

Table 3 First four eigenvalues, $\lambda^{*2} = \omega b^2 \sqrt{\rho/D}$, for free symmetric mode-type vibration of the simply-supported parallelogram plate

$1/\phi$	Mode No.	$1/\phi'$ 0.4	0.6	0.8	1.0	1.2	1.4	1.6
0.4	1	17.30	13.97	12.47	11.66	11.18	10.86	10.65
	2	62.04	44.39	32.67	25.78	21.55	18.79	16.90
	3	82.18	57.99	50.17	46.33	41.26	34.31	29.35
	4	97.63	90.64	68.85	53.13	45.29	43.57	42.56
0.6	1	15.75	13.32	12.12	11.46	11.05	10.78	10.59
	2	51.44	38.22	29.41	23.92	20.39	18.02	16.36
	3	72.58	55.37	48.60	44.25	37.79	31.96	27.64
	4	95.05	76.45	59.75	48.80	44.77	43.22	41.64
0.8	1	14.75	12.86	11.88	11.31	10.95	10.71	10.55
	2	44.17	34.02	27.06	22.52	19.50	17.42	15.94
	3	66.99	53.30	46.84	41.18	35.07	30.11	26.35
	4	84.12	66.08	53.87	47.10	44.39	42.83	40.11
1.0	1	14.07	12.52	11.69	11.19	10.88	10.66	10.51
	2	39.04	31.00	25.31	21.45	18.80	16.94	15.59
	3	63.19	51.30	44.35	38.25	32.90	28.62	25.30
	4	72.51	59.03	50.65	46.35	44.07	42.03	37.90
1.2	1	13.57	12.27	11.54	11.10	10.81	10.62	10.48
	2	35.30	28.74	23.95	20.59	18.23	16.55	15.31
	3	60.01	48.76	41.49	35.80	31.14	27.39	24.43
	4	64.37	54.76	49.11	45.88	43.71	40.35	36.12

Table 4 First four eigenvalues, $\lambda^{*2} = \omega b^2 \sqrt{\rho/D}$, for free antisymmetric mode-type vibration of the simply-supported parallelogram plate

$1/\phi$	Mode No.	$1/\phi'$ 0.4	0.6	0.8	1.0	1.2	1.4	1.6
0.4	1	37.68	26.02	20.18	17.00	15.08	13.85	13.00
	2	48.79	43.80	41.97	37.57	30.45	25.65	22.34
	3	92.49	66.42	49.36	41.50	40.82	40.45	37.49
	4	119.6	83.57	64.35	55.69	50.52	44.38	40.32
0.6	1	31.73	23.32	18.80	16.20	14.58	13.51	12.76
	2	46.99	43.08	40.90	34.12	28.28	24.23	21.35
	3	74.90	56.19	44.09	41.28	40.69	40.02	35.05
	4	105.5	76.72	61.29	53.36	47.49	41.25	40.26
0.8	1	27.91	21.47	17.80	15.60	14.20	13.25	12.57
	2	45.81	42.33	38.04	31.52	26.63	23.12	20.57
	3	63.12	49.35	42.41	41.13	40.53	37.88	33.08
	4	94.84	71.82	58.49	50.59	44.17	40.56	40.22
1.0	1	25.31	20.13	17.05	15.14	13.89	13.04	12.43
	2	44.87	40.98	35.08	29.51	25.31	22.23	19.94
	3	54.98	45.22	41.99	41.01	40.04	35.68	31.46
	4	86.84	67.56	55.36	47.38	41.61	40.45	40.18
1.2	1	23.43	19.12	16.46	14.77	13.65	12.86	12.30
	2	43.89	38.51	32.68	27.91	24.24	21.49	19.41
	3	49.35	43.57	41.77	40.88	38.26	33.84	30.11
	4	78.64	62.73	51.75	44.33	40.84	40.40	40.13

plate analysis, as described herein, has been utilized to obtain these mode shapes. The mode shape of Fig. 8, which corresponds, in fact, to the fundamental mode of vibration has contour lines of the type anticipated. The contour lines of the first antisymmetric type mode of Fig. 9 are also of an anticipated pattern. Because the parallelogram plate is of equal edge length, and therefore of rhombic shape, we anticipate a straight node line joining the closest pair of opposite corners. This is seen to be the case. In an even higher mode of this type a straight node line will join the other pair of opposite corners. Such node lines can be predicted with confidence because in this case the plate of interest happens to have a rhombic shape.

6 Summary and Conclusions

The superposition method utilized with judiciously chosen building blocks is found to constitute a powerful tool for the obtaining of highly accurate analytical-type solutions for free

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

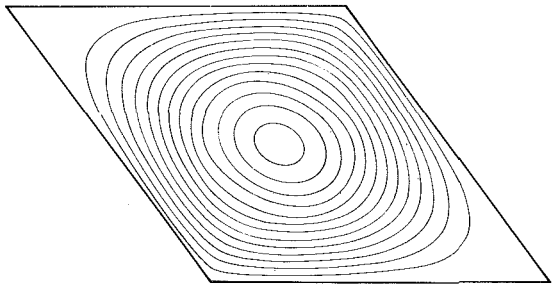


Fig. 8 Computed mode shape for parallelogram plate of equal edge lengths in fundamental mode vibration; $\phi = 4/3$, $\phi' = 4$

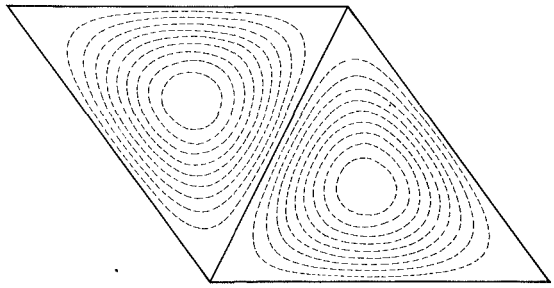


Fig. 9 Computed mode shape for parallelogram plate of Fig. 8 in first antisymmetric-type mode vibration

vibration of simply-supported parallelogram plates. The analysis has been simplified significantly by taking advantage of symmetry which is inherent in this particular problem. In order to take advantage of this symmetry it has been necessary to introduce a modified alternating Lévy-type solution.

It has been possible to conduct numerous verification tests on the computed eigenvalues and mode shapes. This was achieved by comparing eigenvalues for parallelogram plates of equal edge lengths with those obtained for geometrically equivalent rhombic plates. Eigenvalues for the rhombic plates had been obtained earlier by means of a much simpler analysis. Agreement between the two sets of eigenvalues was found to be excellent, thereby adding a high degree of confidence to both analytical procedures and to the tabulated results.

It will be appreciated that there exists several combinations of classical boundary conditions for which the rules of symmetry as discussed herein will apply. An obvious example will be parallelogram plates with all edges clamped. Of course, even if symmetry considerations do not apply the problems will still be amenable to solution by the superposition method but more building blocks will be required resulting in a larger eigenvalue matrix. The obtaining of solutions for these additional problems will form the subject matter of future research.

To the author's knowledge this constitutes the first highly accurate and comprehensive analysis of the free vibration of simply-supported parallelogram plates.

References

Gorman, D. J., 1988, "Accurate Free Vibration Analysis of Rhombic Plates With Simply Supported and Fully Clamped Edge Conditions," *Journal of Sound and Vibration*, Vol. 125, No. 1, pp. 281–290.

Gorman, D. J., 1989, "Accurate Free Vibration Analysis of Right Triangular Plates with one Free Edge," *Journal of Sound and Vibration*, Vol. 131, No. 1, pp. 115–125.

Kaul, P. K., and Cadambe, V., 1956, "The Natural Frequencies of Thin Skew Plates," *Aeron. Quarterly*, Vol. 74, No. 3, pp. 337–352.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Bending Frequency of a Rotating Beam With an Elastically Restrained Root

Sen Yung Lee
Associate Professor,
Mechanical Engineering Department.

Yee Hsiung Kuo
Graduate Student,
Institute of Aeronautics and Astronautics.

National Cheng Kung University,
Tainan, Taiwan 70101
Republic of China

Upper and lower bounds of the fundamental bending frequency of a rotating uniform beam with an elastically restrained root are obtained by the Rayleigh's and minimum principles, respectively. It is shown that the fundamental bending frequency of the rotating uniform beam with rotational flexibility only is always higher than that of the nonrotating beam. If the setting angle is not equal to zero, the fundamental bending frequency of the rotating uniform beam with translational flexibility can be less than that of the nonrotating beam, and the phenomenon of divergence instability (tension buckling due to the centrifugal force) may occur. Finally, the influence of hub radius, setting angle, rotational speed, and elastic root restraints on the fundamental bending frequency of the beam is also investigated numerically by the transfer matrix method.

Introduction

Rotating beams, which have importance in many practical applications such as turbine blades, airplane propellers, and long and flexible rotating space booms, have been studied by numerous investigators (Leissa, 1981; Ramamurti et al. 1984). In the vibrational analysis, the structures were often modeled as cantilever beams vibrating in flexure (Schilhansl, 1958; Lo et al., 1960; Pnueli, 1972, Peters et al., 1980; Hodges, 1981; Storti et al., 1987). The upper and lower bounds of the fundamental bending frequency of a rotating uniform beam with a clamped root were obtained analytically by Schilhansl (1958) and Pnueli (1972). It was concluded that the bending frequency of rotating uniform cantilever beams, because of centrifugal stiffening, is higher than that of nonrotating beams.

In practice, however, the flexibility at the root of rotating beams is not always negligible, especially in bladed disk assemblies with complex fixture like "fir tree" or "pin joint" roots. The actual beam root conditions may be regarded as elastically restrained conditions. The effect of root flexibility on the bending frequencies of rotating beams were investigated by Boyce (1956), Wright et al. (1982), and recently by Abbas (1985). The results indicated that the translational flexibility has little effect on the fundamental bending frequency and an important effect on the bending frequencies of higher modes, while the rotational flexibility has an important effect on the fundamental bending frequency and a little effect on the higher modes. Nevertheless, in the existing literatures, there is still no analytical study on the upper and lower bounds of the bending frequencies of a rotating beam with an elastically restrained root.

In the present study, the Rayleigh's and minimum principles are used to obtain the upper and lower bounds of the fundamental bending frequency of a rotating uniform Bernoulli-Euler beam with an elastically restrained root. The influence of hub radius, setting angle, rotational speed, and elastic root restraints on the fundamental bending frequency of the beam is also investigated numerically by the transfer matrix method (Pestel et al., 1963). Several new physical observations are presented.

Analysis

Consider the free transverse vibrations in the Z-direction of an elastic uniform Bernoulli-Euler beam elastically restrained mounted with setting angle θ on a hub with radius r_o rotating with constant angular velocity Ω, as shown in Fig. 1. Here the Coriolis forces are neglected and the mass axis, centroidal axis, and elastic axis are coincident. For time harmonic vibration with angular frequency ω, the flexural displacement $W(X)$ satisfy the following characteristic differential equation:

$$EI\frac{d^4W}{dX^4}-\frac{d}{dX}\left(N\frac{dW}{dX}\right)-m(\omega^2+\Omega^2\sin^2\theta)W=0,\quad X\in(0,l) \tag{1}$$

where l is the length of the beam and N is the centrifugal force

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, OH, June 16–19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 30, 1989; final revision, Mar. 19, 1990.

Paper No. 91-APM-7.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$N(X)=\Omega^2\int_X^l m(s+r_o)ds. \tag{2}$$

The associated boundary conditions are

$$\text{at } X=0:\ EI\frac{d^2W}{dX^2}=K_\theta\frac{dW}{dX},\qquad EI\frac{d^3W}{dX^3}-N\frac{dW}{dX}=-K_TW \tag{3}$$

and

$$\text{at } X=l:\ \frac{d^2W}{dX^2}=0,\qquad \frac{d^3W}{dX^3}=0, \tag{4}$$

where E, I, and m denote Young's modulus, moment of inertia, and mass per unit length of beam, respectively. K_θ and K_T are rotational and translational spring stiffness at $X = 0$, respectively.

In terms of the following nondimensional quantities

$$x=\frac{X}{l},\quad V=\frac{W}{l},\quad R=\frac{r_o}{l},\quad \Lambda^2=\frac{m\omega^2l^4}{EI},\quad \alpha^2=\frac{m\Omega^2l^4}{EI},\quad \gamma^2=\Lambda^2+\alpha^2\sin^2\theta,\quad \beta_\theta=\frac{EI}{K_\theta l},\quad \beta_T=\frac{EI}{K_Tl^3}, \tag{5}$$

the governing characteristic differential equation can be rewritten in the following nondimensional form:

$$V''''-\alpha^2[\eta V']'-\gamma^2V=0,\quad x\in(0,\ 1) \tag{6}$$

where

$$\eta=R(1-x)+\frac{1}{2}(1-x^2). \tag{7}$$

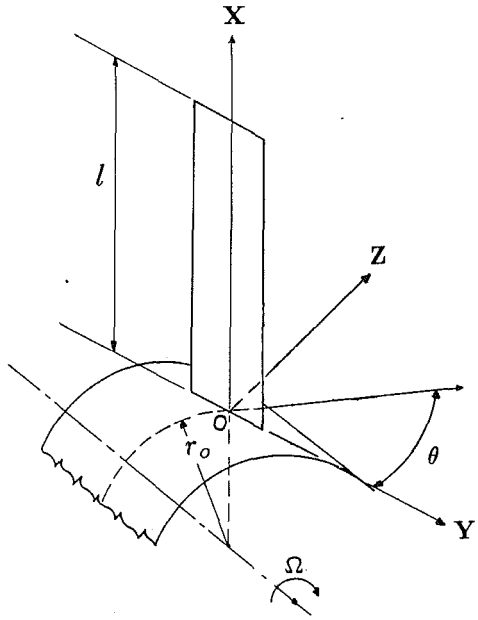


Fig. 1 Beam on a rotating disk as a model of a turbomachinery blade

The associated boundary conditions become

$$\text{at } x=0:\ \beta_\theta V''-V'=0,\qquad \beta_TV'''-\alpha^2\left(R+\frac{1}{2}\right)\beta_TV'+V=0, \tag{8}$$

and

$$\text{at } x=1:\ V''=0,\qquad V'''=0, \tag{9}$$

where the prime system refers to differentiation with respect to x.

Upper Bound. The square of upper bound of the fundamental bending frequency can be obtained by applying the Rayleigh's principle. The Rayleigh's quotient of the system is

$$\gamma^2=\left\{\int_0^1[(V'')^2+\alpha^2\eta(V')^2]dx+\frac{1}{\beta_\theta}[V'(0)]^2+\frac{1}{\beta_T}V^2(0)\right\}\Big/\int_0^1V^2dx. \tag{10}$$

By substituting the following admissible function

$$V(x)=12\alpha^2\left(R+\frac{1}{2}\right)\beta_\theta\beta_T+24\beta_T+12\beta_\theta x+6x^2-4x^3+x^4, \tag{11}$$

which is the static deformation of uniform beam under uniformly distributed transverse loads, satisfying the prescribed boundary conditions, into the Rayleigh's quotient, it yields

$$\gamma_u^2=\frac{\left\{\begin{array}{c}28.8+144\beta_\theta+576\beta_T+144\alpha^4\left(R+\frac{1}{2}\right)^2\beta_\theta^2\beta_T\\+\alpha^2[(2.71+20.8\beta_\theta+48\beta_\theta^2+288\beta_\theta\beta_T)+R(3.6+28.8\beta_\theta+72\beta_\theta^2+576\beta_\theta\beta_T)]\end{array}\right\}}{\left\{\begin{array}{c}2.31+20.8\beta_\theta+57.6\beta_T+48\beta_\theta^2+576\beta_T^2+288\beta_\theta\beta_T\\+\alpha^2\left(R+\frac{1}{2}\right)(28.8\beta_\theta\beta_T+576\beta_\theta\beta_T^2+144\beta_\theta^2\beta_T)+144\alpha^4\left(R+\frac{1}{2}\right)^2\beta_\theta^2\beta_T^2\end{array}\right\}}. \tag{12}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Fundamental bending frequency of a nonrotating uniform beam

Error (%) = $\frac{\Lambda_u - \Lambda_o}{\Lambda_o} \times 100(\%)$

[Λ_o is the exact solution given in Afolabi (1986)]

$\beta_T = 0$				$\beta_\theta = 0$			
β_θ	Λ_u	Λ_o	Error (%)	β_T	Λ_u	Λ_o	Error (%)
1.	1.5588	1.5573	0.10	1.	0.9752	0.9752	0.
0.1	2.9780	2.9678	0.34	0.1	2.4994	2.4872	0.49
0.01	3.4614	3.4477	0.40	0.01	3.4258	3.3843	1.23
0.001	3.5231	3.5090	0.40	0.001	3.5212	3.5027	0.53
0.	3.5301	3.5160	0.40	0.	3.5301	3.5160	0.40

The square of the upper bound of the fundamental bending frequency is thus

$$\Lambda_u^2 = (\gamma_u^2 - \alpha^2 \sin^2\theta), \tag{13}$$

which is a function of the ratio of the hub radius to the beam length, bending rigidity, rotational speed, setting angle, and two elastic restraint constants.

For a nonrotating beam, $\alpha = 0$. Equation (12) becomes

$$\Lambda_u^2 = \frac{28.8 + 144\beta_\theta + 576\beta_T}{[2.31 + 20.8\beta_\theta + 57.6\beta_T + 48\beta_\theta^2 + 576\beta_T^2 + 288\beta_\theta\beta_T]}. \tag{14}$$

In Table 1, the fundamental bending frequency of a nonrotating beam with various rotational and translational flexibilities, determined from equation (14), is compared with the exact solution obtained by Afolabi (1986). It shows the difference is less than 1.3 percent.

Lower Bound. A lower bound is obtained by the method adopted from Boyce (1956) and Pnueli (1972). From equation (10), the variational formula of the problem is

$$\gamma^2 = \min\left\{\left[\int_0^1 \{(V'')^2 + \alpha^2\eta(V')^2\right]dx + \frac{1}{\beta_\theta}[V'(0)]^2 + \frac{1}{\beta_T}V^2(0)\right] \Big/ \int_0^1 V^2 dx\right\}. \tag{15}$$

Denote

$$\gamma_0^2 = \min\left\{\left[\int_0^1 (V'')^2 dx + \frac{1}{\beta_\theta}[V'(0)]^2 + \frac{1}{\beta_T}V^2(0)\right] \Big/ \int_0^1 V^2 dx\right\}. \tag{16}$$

$$\gamma_1^2 = \min\left\{\int_0^1 \alpha^2\eta(V')^2 dx \Big/ \int_0^1 V^2 dx\right\}. \tag{17}$$

Clearly,

$$\gamma^2 \geq \gamma_0^2 + \gamma_1^2. \tag{18}$$

The solution of the variational problem of equation (16) yields just the square of fundamental bending frequency with no rotation. Hence, only the solution of equation (17) requires some further considerations.

By equations (7), (10), and (17)

$$\gamma_1^2 = \min\left\{\alpha^2\int_0^1 \left[\frac{1}{2}(1-x^2)(V')^2 + R(1-x)(V')^2\right]dx \Big/ \int_0^1 V^2 dx\right\}$$
$$= \alpha^2\int_0^1\left[\frac{1}{2}(1-x^2)(V_*')^2 + R(1-x)(V_*')^2\right]dx \Big/ \int_0^1 V_*^2 dx, \tag{19}$$

i.e., V_* is the solution of this variational problem. The domain of integration is $0 \leq x \leq 1$ and in this domain $\frac{1}{2}(1+x) \leq 1$. Because all integrands in equation (19) are non-negative, this yields

$$\gamma_1^2 \geq \alpha^2(1+R)\int_0^1 \frac{1}{2}(1-x^2)(V_*')^2 dx \Big/ \int_0^1 V_*^2\, dx$$
$$\geq \alpha^2(1+R)\ \min\left\{\int_0^1 \frac{1}{2}(1-x^2)(V')^2 dx \Big/ \int_0^1 V^2\, dx\right\}. \tag{20}$$

Denote

$$\gamma_2^2 = \int_0^1 \frac{1}{2}(1-x^2)(V')^2 dx \Big/ \int_0^1 V^2 dx. \tag{21}$$

The minimum value of γ_2^2 can be obtained by assuming the beam is in rigid body motion. Consequently, from equation (5), (18), (20), and (21), the square of lower bound of the fundamental frequency is

$$\Lambda_L^2 = \gamma_0^2 + \alpha^2[(1+R)\gamma_2^2 - \sin^2\theta]. \tag{22}$$

Rotating Beam With Rotational Flexibility Only. For a rotating beam with rotational flexibility only at the root of the beam, the square of the upper bound of the fundamental bending frequency can be obtained by setting $\beta_T = 0$ in equations (12) and (13)

$$\Lambda_u^2 = \frac{28.8 + 144\beta_\theta + \alpha^2\psi}{23.1 + 20.8\beta_\theta + 48\beta_\theta^2}, \tag{23}$$

where

$$\psi(R,\theta,\beta_\theta) = [2.71 + 20.8\beta_\theta + 48\beta_\theta^2] + R[3.6 + 28.8\beta_\theta + 72\beta_\theta^2] - \sin^2\theta[2.31 + 20.8\beta_\theta + 48\beta_\theta^2]. \tag{24}$$

In terms of the fundamental bending frequency of a nonrotating beam Λ_o, equation (23) becomes

$$\Lambda_u^2 = \Lambda_o^2 + \alpha^2[\psi^* - \sin^2\theta] \tag{25}$$

where

$$\psi^* = \frac{[2.71 + 20.8\beta_\theta + 48\beta_\theta^2] + R[3.6 + 28.8\beta_\theta + 72\beta_\theta^2]}{[2.31 + 20.8\beta_\theta + 48\beta_\theta^2]}. \tag{26}$$

The square of lower bound of the fundamental bending frequency is obtained by assuming the beam is in rigid body motion, rotating about the root $x = 0$,

$$V(x) = Cx. \tag{27}$$

Substituting equation (27) into equation (21), one obtains $\gamma_2^2 = 1$. Consequently, it follows from equation (22) that the square of lower bound of the fundamental bending frequency is

$$\Lambda_L^2 = \Lambda_o^2 + \alpha^2[1 + R - \sin^2\theta]. \tag{28}$$

Combining the square of upper and lower bounds of the fundamental bending frequency, equations (25) and (28), one reaches

$$\Lambda_o^2 + \alpha^2[\psi^* - \sin^2\theta] \geq \Lambda^2 \geq \Lambda_o^2 + \alpha^2[1 + R - \sin^2\theta], \tag{29}$$

where ψ^* is defined in equation (26).

If the rotational flexibility of the system is zero, then by setting $\beta_\theta = 0$ in equation (29), the square of upper and lower bounds of the fundamental bending frequency of a rotating cantilever beam is obtained

$$\Lambda_o^2 + \alpha^2[1.173 + 1.558R - \sin^2\theta] \geq \Lambda^2 \geq \Lambda_o^2 + \alpha^2[1 + R - \sin^2\theta], \tag{30}$$

which is the same as that given by Puneli (1972). Since $1 + R - \sin^2\theta \geq 0$ for $R \geq 0$, it confirms the conclusion given by Schilhansl (1958) who concluded that the bending frequency of rotating uniform cantilever beams, because of centrifugal stiffening, is higher than that of nonrotating beams.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

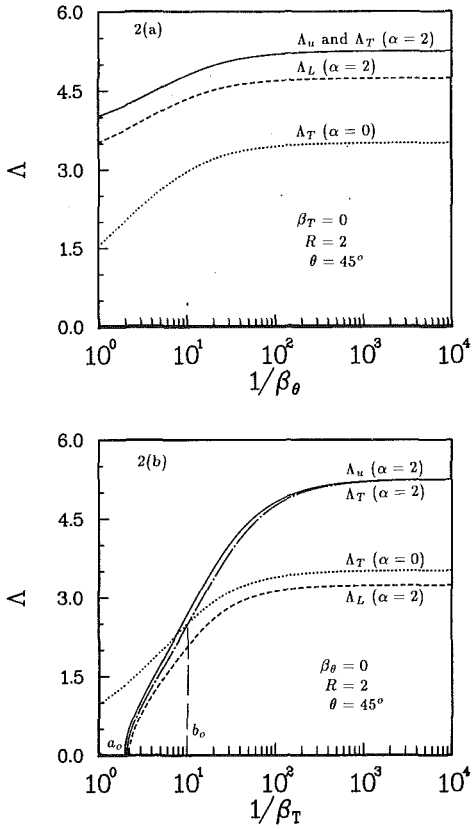


Fig. 2 Fundamental bending frequency of rotating and nonrotating beams with an elastically restrained root; 2(*a*) with rotational flexibility only, 2(*b*) with translation flexibility only.

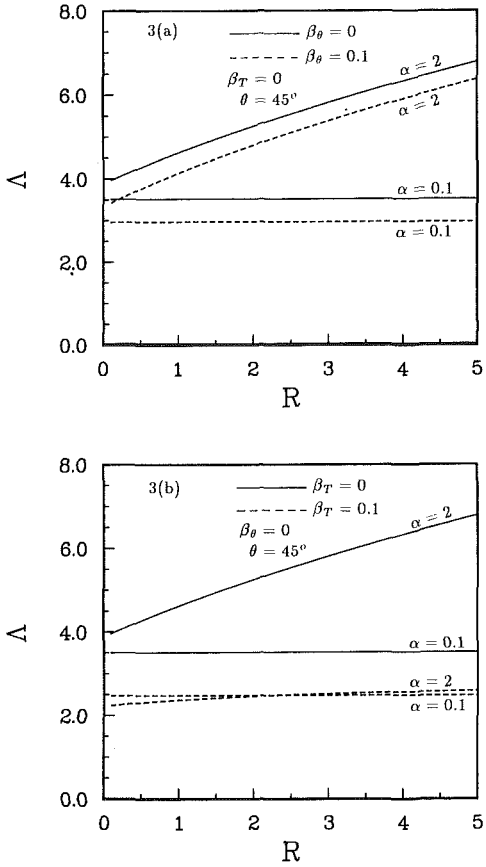


Fig. 3 Influence of hub radius on the fundamental bending frequency of a rotating beam with an elastically restrained root (evaluated by the transfer matrix method); 3(*a*) with rotational flexibility only, 3(*b*) with translational flexibility only.

Because the term $1 + R - \sin^2\theta$ also appears in equation (29), one can generalize the conclusion given by Schilhansl (1958) and conclude that the fundamental bending frequency of a rotating beam with rotational flexibility only is always greater than that of nonrotating beam.

Rotating Beam With Translational Flexibility Only. By setting $\beta_\theta = 0$ in equation (12), the square of upper bound of the fundamental bending frequency of a rotating beam with translational flexibility only, is reduced to the following form

$$\Lambda_u^2 = \frac{28.8 + 576\beta_T + \alpha^2\phi}{2.31 + 57.6\beta_T + 576\beta_T^2}, \tag{31}$$

where

$$\phi(R,\theta,\beta_T) = 2.71 + 3.6R - \sin^2\theta(2.31 + 57.6\beta_T + 576\beta_T^2). \tag{32}$$

In terms of the fundamental bending frequency of a nonrotating beam system, equation (31) becomes

$$\Lambda_u^2 = \Lambda_o^2 + \alpha^2(\phi^* - \sin^2\theta), \tag{33}$$

where

$$\phi^* = (2.71 + 3.6R)/(2.31 + 57.6\beta_T + 576\beta_T^2). \tag{34}$$

The square of lower bound of the fundamental bending frequency is obtained by assuming the beam is in rigid body translation motion

$$V(x) = C. \tag{35}$$

Substituting equation (35) into (21), it yields $\gamma_2^2 = 0$. It follows from equations (22) that the square of the lower bound of the fundamental bending frequency is

$$\Lambda_L^2 = \Lambda_o^2 - \alpha^2\sin^2\theta. \tag{36}$$

Combining the square of upper and lower bounds of the fundamental bending frequency, equations (33) and (36), yields

$$\Lambda_o^2 + \alpha^2[\phi^* - \sin^2\theta] \geq \Lambda^2 \geq \Lambda_o^2 - \alpha^2\sin^2\theta, \tag{37}$$

where ϕ^* is defined in equation (34).

From equations (33) and (34), it is observed that if the setting angle is not equal to zero and the translational flexibility is relatively high, i.e., β_T is large, then $\phi^* - \sin^2\theta$ can be negative. In this case, the fundamental bending frequency of rotating beam can be less than that of nonrotating beam. Especially, if the rotational speed of the beam is sufficiently large, the square of the upper bound of the fundamental bending frequency Λ_u^2 can be less than zero. In this case, the fundamental bending frequency of the rotating beam is pure imaginary and the phenomenon of divergence instability (tension buckling due to centrifugal force) occurs. One can consider the case when the frequency goes to zero as a buckling boundary of the system.

Numerical Results and Discussion

To illustrate the previous analysis and investigate the influence of hub radius, setting angle, and rotational speed on the fundamental bending frequency of a rotating beam with an elastically restrained root, several numerical results are presented and discussed.

In Fig. 2, the upper bound (Λ_u) and the lower bound (Λ_L) of the fundamental bending frequency of the rotating beam and the numerical result (λ_T obtained by the transfer matrix method (Pestel et al. (1963)) for the fundamental bending frequency of rotating and the nonrotating beams are shown. It can be found that the upper bound and the numerical result

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

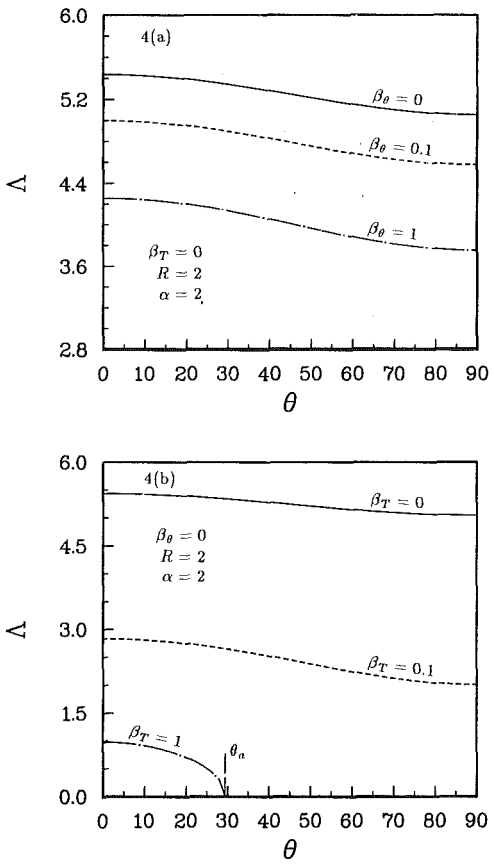


Fig. 4 Influence of setting angle on the fundamental bending frequency of a rotating beam with an elastically restrained root (evaluated by transfer matrix method), 4(*a*) with rotational flexibility only, 4(*b*) with translational flexibility only.

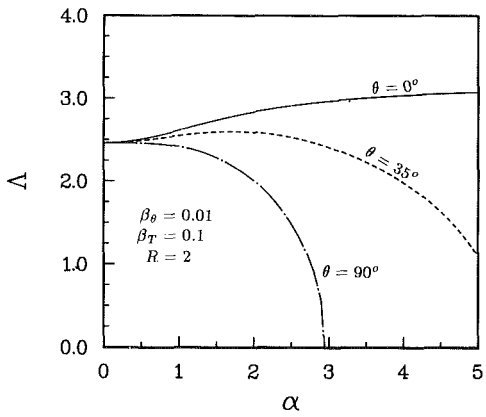


Fig. 5 Influence of rotational speed on the fundamental bending frequency of a rotating beam with an elastically restrained root and different setting angles (evaluated by transfer matrix method)

are almost coincident. The fundamental bending frequency of a rotating beam with rotational flexibility only is always greater than that of a nonrotating beam. However, only when $\frac{1}{\beta_T}$ is greater than b_o the fundamental bending frequency of a rotating beam with translational flexibility only is greater than that of a nonrotating beam. Especially, when $\frac{1}{\beta_T}$ is less than a_o, which is a buckling boundary of the system, the phenomenon of divergence instability occurs. These numerical results confirm the conclusions in our previous analysis. The increase of $\frac{1}{\beta_T}$ and $\frac{1}{\beta_\theta}$ will increase the fundamental bending frequency of rotating and nonrotating beams. The influence of rotational flexibility on the fundamental bending frequency of the nonrotating beam is greater than that of the rotating beam. However, the influence of the translational flexibility on the fundamental bending frequency of the rotational beam is more pronounced than that of the nonrotating beam.

Figure 3 illustrates the influence of hub radius on the fundamental bending frequency of a rotating beam with an elastically restrained root. It is indicated that the relation between the fundamental bending frequency and hub radius of the beam is almost linear. When the hub radius increases, the fundamental bending frequency increases. However, at low rotational speed, the influence of hub radius on the fundamental bending frequency is almost negligible. This conclusion is similar to that of the rotating cantilever beam, given by Boyce (1956).

From Fig. 4, one observes that when the setting angle increases, the fundamental bending frequency of the beam decreases. Around the buckling boundary point θ_a, the influence of the setting angle on the fundamental bending frequency is very significant.

Figure 5 shows the influence of rotational speed on the fundamental bending frequency of a beam with an elastically restrained root and different setting angles. For the beam with small setting angle, the increase of rotational speed may increase the fundamental bending frequency of the beam. However, if the setting angle is large, the increase of rotational speed may decrease the fundamental bending frequency.

Conclusions

In this paper, the upper and lower bounds of the fundamental bending frequency of a rotating uniform beam with an elastically restrained root are obtained by the Rayleigh's and minimum principle, respectively. The influence of hub radius, setting angle, rotational speed, and elastic root restraints on the fundamental bending frequency of the beam is investigated numerically by the transfer matrix method. It is shown that:

(1) The fundamental bending frequency of the rotating uniform beam with rotational flexibility only is always higher than that of the nonrotating beam.

(2) If setting angle is not equal to zero, the fundamental bending frequency of the rotating uniform beam with translational flexibility can be less than that of the nonrotating beam and the phenomenon of divergence instability (tension buckling due to the centrifugal force) may occur.

(3) When $\frac{1}{\beta_\theta}$ or $\frac{1}{\beta_T}$ increases, the fundamental bending frequency increases. The influence of rotational flexibility on the fundamental bending frequency of the nonrotating beam is greater than that of the rotating beam. However, the influence of translational flexibility on the frequency of the nonrotating beam is less than that of the rotating beam.

(4) The relation between the fundamental bending frequency and hub radius of the beam is almost linear. When the hub radius increases, the frequency increases. However, at low rotational speed, the influence of hub radius on the fundamental bending frequency is negligible.

(5) When the setting angle increases, the fundamental bending frequency of the beam decreases. The influence of the setting angle on the fundamental bending frequency is very significant when the phenomenon of divergence instability is about happened.

(6) For a beam with an elastically restrained root and small setting angle, the increase of rotational speed may increase the fundamental bending frequency of the beam. However, if the setting angle is large, the increase of rotational speed may decrease the fundamental bending frequency.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

References

Abbas, B. A. H., 1985, "Dynamic Analysis of Thick Rotating Blades with Flexible Roots," *The Aeronautical Journal*, Vol. 89, pp. 10–16.

Afolabi, D., 1986, "Natural Frequencies of Cantilever Blades with Resilent Roots," *Journal of Sound and Vibration*, Vol. 110, pp. 429–441.

Boyce, W. E., 1956, "Effect of Hub Radius on the Vibrations of a Uniform Bar," ASME JOURNAL OF APPLIED MECHANICS, Vol. 23, pp. 287–290.

Hodges, D. H., 1981, "An Approximate Formula for the Fundamental Natural Frequency of a Uniform Rotating Beam Clamped off the Axis of Rotation," *Journal of Sound and Vibration*, Vol. 77, pp. 11–18.

Leissa, A., 1981, "Vibrational Aspects of Rotating Turbomachinery Blades," *ASME Applied Mechanics Reviews*, Vol. 34, pp 629–635.

Lo, H., Goldberg, J. E., and Bogdanoff, J. L., 1960, "Effect of Small Hub-Radius Change on Bending Frequencies of a Rotating Beam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 27, pp. 548–550.

Pestel, E. C. and Leckie, F. A., 1963, *Matrix Methods in Elastomechanics*, McGraw-Hill, New York.

Peters, D. A., and Hodges, D. H., 1980, "In-Plane Vibration and Buckling of a Rotating Beam Clamped Off the Axis of Rotation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 398–402.

Pnueli, D., 1972, "Natural Bending Frequency Comparable to Rotational Frequency in Rotating Cantilever Beam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 602–604.

Ramamurti, V., and Balasubramanian, P., 1984, "Analysis of Turbomchinery Blades—A Review," *The Shock and Vibration Digest*, Vol. 16, pp. 13–28.

Schilhansl, M. J., 1958, "Bending Frequency of a Rotating Cantilever Beam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 28–30.

Storti, D., and Aboelnaga, Y., 1987, "Bending Vibrations of a Class of Rotating Beams with Hypergeometric Solutions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 311–314.

Wright, A. D., Smith, C. E., Thresher, R. W., and Wang, J. L. C., 1982, "Vibration Modes of Centrifugally Stiffened Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, 197–202.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Dynamics of Multibody Systems With Known Configuration Changes

J. J. McPhee
Lecturer.

R. N. Dubey
Professor,
Mem. ASME

Department of Mechanical Engineering,
University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1

The equations of motion are derived for a system with inertial properties that are varying in time as a result of known relative motions between the rigid bodies comprising the system. This vector-dyadic formulation has been encoded into a computer program, making use of the conformal rotation vector for the representation of rotations. The numerical simulation of two different physical systems is presented in order to illustrate the dynamic effects of the changing inertial properties, and the usefulness of the encoded formulation for modeling these effects.

1 Introduction

In the first two decades following World War II, a flurry of papers dealing with the dynamics of systems with time-varying inertial properties (mostly rockets) appeared in the open literature (see Grubin, 1963, for further references). After two more decades of relative disinterest, recent papers have appeared by authors re-examining the theory governing such systems (eg., Ge, 1984; Kapoulitsas, 1987), and interested in its application to space structure deployment simulation (Keat and Turner, 1987). By definition, the term "inertial properties" is taken to refer to not only the total mass of a system, but also the distribution of this mass. As discussed by McPhee (1989), one mechanism by which the inertial properties may vary with time is a change in configuration, which obviously alters the distribution of mass within the system. The main goal of this current work is to illustrate the effects of *known configuration changes* on the three-dimensional motion of a material system.

Such a system has been modeled in this paper as an unconstrained rigid base vehicle that contains one or more moving rigid bodies. The motions of these auxiliary bodies relative to the base vehicle are specified functions of time, and are responsible for the changing configuration of the overall system. This particular multibody model has been employed previously by Roberson (1958), and selected for this current work because of its ability to represent a wide variety of physical systems. Furthermore, the equations of motion for this model are mathematically tractable and can be expressed in a form that explicitly reveals the dynamic effects of the changing configuration.

Roberson was specifically interested in determining the effects of internal moving parts on a satellite vehicle. Instead of solving for the motion of the satellite, Roberson derived an "attitude-perturbation torque" as a function of the internal motions. Unfortunately, the usefulness of Roberson's torque term to this present investigation is limited by the fact that it is also a function of the unknown angular motion of the base vehicle. In a later analysis of the same multibody system by Grubin (1962), a procedure for explicitly determining the angular acceleration of the base mass is described, but the actual calculations are not carried out to completion.

In this paper, the equations of motion are derived in a direct manner using vector-dyadic notation, and solved for both the angular and translational accelerations of the base vehicle. By expressing the equations in this final form, the influence of the moving secondary bodies on the motion of the primary vehicle is made explicit. The current work further extends the analyses of Roberson and Grubin by providing an inverse dynamic solution for the internal forces and moments between the base mass and each auxiliary body. These internal loads can be calculated once the motion of the entire system has been determined, and are of importance because they provide the desired relative motions of the secondary bodies.

The final vector-dyadic equations of motion are then resolved into scalar components, resulting in a set of 12 first-order ordinary differential equations governing the unconstrained motion of the base vehicle. Briefly described is a Fortran-77 computer program which has been developed to numerically integrate these highly nonlinear equations using a fourth-order Runge-Kutta algorithm. Once the parameters for a specific system have been input to this program, it will numerically simulate the motion of the system and, if requested, perform the incorporated inverse dynamic solution for the internal forces and moments. Rotations are represented in the program by the conformal rotation vector proposed by Milenkovic (1982), a set of three parameters that are quite attractive from a computational perspective. Finally, two examples are presented which illustrate the utility of the en-

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 13, 1989; final revision, Oct. 30, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

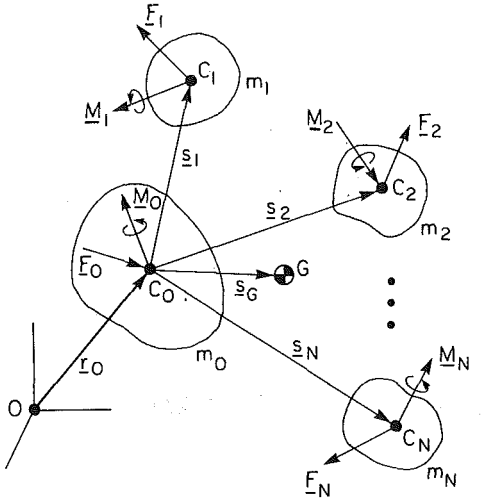


Fig. 1 Multibody model of system

coded formulation and its applicability to various physical systems.

2 System Model

The multibody model employed in this paper to represent a physical system with a time-varying configuration is illustrated in Fig. 1. The rigid base vehicle has a mass m_o and an absolute angular velocity ω_o, and its center of mass C_o is located in inertial space by the position vector $\mathbf{r}_o$. In turn, each auxiliary body has a mass m_i ($i = 1, \ldots, N$), an angular velocity relative to the base vehicle $\omega_{i/o}$, and a relative position vector $\mathbf{s}_i$. By definition, the relative position of the center of mass of the entire system is given by:

$$\mathbf{s}_G = \frac{1}{m}\sum_{i=1}^{N} m_i \mathbf{s}_i, \tag{1}$$

and in general, $\mathbf{s}_G$ varies with time as the center of mass G moves around within the system. Finally, the external loading on the system is modeled by an applied force $\mathbf{F}_i$ and moment $\mathbf{M}_i$ acting at the center of mass of each body in the system.

It will be assumed that the relative position $\mathbf{s}_i$ and angular velocity $\omega_{i/o}$ of each secondary body are specified functions of time. This means that the system has six degrees-of-freedom corresponding to the unconstrained rotational and translational motion of the primary vehicle. In order to derive the governing equations of motion, it is also assumed that for each body in the system the applied forces ($\mathbf{F}_i$) and moments ($\mathbf{M}_i$) will be specified, as well as the mass (m_i) and inertia dyadic with respect to the center of mass (I_i).

3 Equations of Motion

3.1 Principle of Linear Momentum. Since the total system mass $m = \Sigma_{i=0}^{N} m_i$ is constant, the principle of linear momentum reduces to:

$$\mathbf{F} = m\mathbf{a}_G \tag{2}$$

where $\mathbf{a}_G$ is the absolute translational acceleration of G and $\mathbf{F}$ is the resultant external force given by:

$$\mathbf{F} = \sum_{i=0}^{N} \mathbf{F}_i \tag{3}$$

Equation (2) reveals that the translational motion of G is unaffected by the motion of secondary bodies unless the resultant force $\mathbf{F}$ is configuration dependent. As pointed out in elementary dynamics texts, the translational motion of the system's center of mass would not be changed if the system were replaced by a single particle of mass m acted on by a force $\mathbf{F}$.

Using equation (3) and the relationship

$$\mathbf{a}_G = \mathbf{a}_o + \ddot{\mathbf{s}}_G \tag{4}$$

where $\mathbf{a}_o$ is the acceleration of C_o and the dot ($\cdot$) notation represents a time derivative measured by an observer in inertial space, equation (2) can be solved explicitly for $\mathbf{a}_o$:

$$\mathbf{a}_o = \frac{1}{m}\sum_{i=0}^{N} \mathbf{F}_i - \ddot{\mathbf{s}}_G. \tag{5}$$

The relative acceleration $\ddot{\mathbf{s}}_G$ is obtained by twice differentiating equation (1):

$$\ddot{\mathbf{s}}_G = \frac{1}{m}\sum_{i=1}^{N} m_i \ddot{\mathbf{s}}_i, \tag{6}$$

and in turn, $\ddot{\mathbf{s}}_i$ is obtained using vector calculus:

$$\ddot{\mathbf{s}}_i = \mathring{\mathring{\mathbf{s}}}_i + \mathbf{a}_o \times \mathbf{s}_i + \omega_o \times (\omega_o \times \mathbf{s}_i) + 2\omega_o \times \mathring{\mathbf{s}}_i. \tag{7}$$

The circle ($\circ$) notation represents a time derivative measured by an observer rotating with the base vehicle, and ω_o and α_o are the angular velocity and acceleration of this base mass, respectively. Since the relative position vector $\mathbf{s}_i$ is a given function of time, $\mathring{\mathbf{s}}_i$ and $\mathring{\mathring{\mathbf{s}}}_i$ are also known functions of time. Equation (5) indicates that the translational acceleration of the base vehicle, unlike $\mathbf{a}_G$, is very much affected by the relative motions of the auxiliary bodies. Equation (7) shows that in order to solve equation (5) for $\mathbf{a}_o$, it will be necessary to obtain the angular motion (ω_o, α_o) of the base vehicle.

3.2 Principle of Angular Momentum. By taking moments about G, the principle of angular momentum takes the simple form:

$$\mathbf{M}_G = \dot{\mathbf{h}}_G \tag{8}$$

where $\mathbf{h}_G$ is the angular momentum of the system about G and $\mathbf{M}_G$ is the resultant external moment about G given by:

$$\mathbf{M}_G = \sum_{i=0}^{N} \mathbf{M}_i + \sum_{i=1}^{N} \mathbf{s}_i \times \mathbf{F}_i - \mathbf{s}_G \times \mathbf{F}. \tag{9}$$

The angular momentum of m_i about its own center of mass is $\mathbf{h}_i$, where

$$\mathbf{h}_i = I_i \cdot \omega_i \tag{10}$$

and $\omega_i = \omega_o + \omega_{i/o}$. Using the translation theorem given by Meirovitch (1970), the angular momentum of m_i about G is:

$$\mathbf{h}_{G_i} = I_i \cdot \omega_i + (\mathbf{s}_i - \mathbf{s}_G) \times m_i(\dot{\mathbf{s}}_i - \dot{\mathbf{s}}_G). \tag{11}$$

Summing up the contribution of all masses including the base vehicle,

$$\mathbf{h}_G = I_o \cdot \omega_o + \mathbf{s}_G \times m_o \dot{\mathbf{s}}_G + \sum_{i=1}^{N} \mathbf{h}_{G_i}. \tag{12}$$

By differentiating equation (12), substituting into equation (8) and manipulating, one obtains:

$$I \cdot \alpha_o + \omega_o \times (I \cdot \omega_o) = \mathbf{M}_G + \Big\{ \mathbf{s}_G \times m[\mathring{\mathring{\mathbf{s}}}_G + 2\omega_o \times \dot{\mathbf{s}}_G] - \sum_{i=1}^{N} [I_i \cdot \dot{\omega}_{i/o} + \omega_i \times (I_i \cdot \omega_i) - \omega_o \times (I_i \cdot \omega_o)] - \sum_{i=1}^{N} [\mathbf{s}_i \times m_i(\mathring{\mathring{\mathbf{s}}}_i + 2\omega_o \times \dot{\mathbf{s}}_i)] \Big\} \tag{13}$$

where $\dot{\omega}_{i/o} = \mathring{\omega}_{i/o} + \omega_o \times \omega_{i/o}$, from vector calculus. Since $\omega_{i/o}$ is a specified function of time, $\mathring{\omega}_{i/o}$ is also a known function

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

of time. Equation (13) thus provides the means for calculating the angular acceleration (α_o) of the base vehicle as a function of known variables and the angular velocity ω_o.

The inertia dyadic I in equation (13) is a short form for:

$$I = I_o + \sum_{i=1}^{N} [I_i + m_i[(\mathbf{s}_i \cdot \mathbf{s}_i)I - \mathbf{s}_i\mathbf{s}_i]] - \left\{ m[(\mathbf{s}_G \cdot \mathbf{s}_G)I - \mathbf{s}_G\mathbf{s}_G] \right\} \quad (14)$$

in which I is the unit dyadic. The first two terms in equation (14) represent the inertia dyadic of the system about C_o, and the last term (in braces) accounts for a shift of reference point from C_o to G. Thus, I is exactly the time-varying inertia dyadic of the entire system about its moving center of mass G. If there are no auxiliary bodies connected to the vehicle, I reduces to I_o, G coincides with C_o, and equation (13) correctly reduces to the standard form for a single rigid body because the three terms in braces drop out of that equation. When auxiliary bodies are present, it is these three terms in equation (13) that account for the time-varying configuration of the system. The first of these terms accounts for the shifting center of mass G, the second term represents the total effect of the relative rotational motion of each auxiliary body, and the last term represents the summed effect of the relative translational motion of each secondary mass.

3.3 Inverse Dynamic Anaysis. Once equations (5) and (13) have been solved for the motion of the base vehicle, the motion of the entire system is known. Assuming that this has been accomplished, an inverse dynamic analysis can be used to determine the internal reaction forces and moments between the primary vehicle and the secondary masses. It is these internal forces and moments that are responsible for the controlled relative motions of the auxiliary masses.

To begin the analysis, a free-body diagram of mass m_i is constructed as shown in Fig. 2. The reaction force and moment on the body have been represented by $\mathbf{P}_i$ and $\mathbf{N}_i$, respectively. In addition to $\mathbf{F}_i$ and $\mathbf{M}_i$, it is assumed that the user has specified $\mathbf{q}_i$, the position of the contact point relative to the center of mass C_i. Newton's second law is used to solve for the internal force $\mathbf{P}_i$:

$$\mathbf{P}_i = m_i\mathbf{a}_i - \mathbf{F}_i \quad (15)$$

in which $\mathbf{a}_i$, the acceleration of C_i, is calculated using:

$$\mathbf{a}_i = \mathbf{a}_o + \ddot{\mathbf{s}}_i \quad (16)$$

in conjunction with equation (7) for $\ddot{\mathbf{s}}_i$.

The principle of angular momentum is applied to obtain the internal moment:

$$\mathbf{N}_i = \dot{\mathbf{h}}_i - \mathbf{M}_i - \mathbf{q}_i \times \mathbf{P}_i \quad (17)$$

in which the absolute time derivative of the angular momentum about C_i is given by:

$$\dot{\mathbf{h}}_i = I_i \cdot \dot{\omega}_i + \omega_i \times (I_i \cdot \omega_i) \quad (18)$$

and $\dot{\omega}_i = \alpha_o + \dot{\omega}_{i/o}$. Using equations (15) and (17), the internal force $\mathbf{P}_i$ and moment $\mathbf{N}_i$ can be calculated from the known motion of the system and the given applied force $\mathbf{F}_i$ and moment $\mathbf{M}_i$.

4 Introduction of Reference Frames

The vector-dyadic notation employed in the previous section is useful for presenting the equations in a concise invariant form. In order to obtain numerical results, however, the components of these vectors and dyadics must be written out explicitly—a process one author (DeBra, 1963) refers to as "coordinatization." To accomplish this, a Cartesian reference frame is defined which has an origin at O and XYZ axes fixed in inertial space. In addition, a moving Cartesian reference

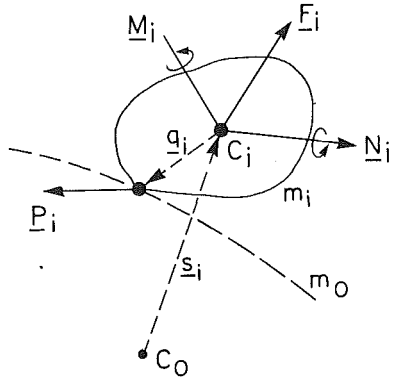


Fig. 2 Free-body diagram of mass m_i

frame is rigidly attached to each body in the system. The axes are designated by $x_iy_iz_i$ ($i = 0, \ldots, N$) and the origins coincide with the centers of mass C_i. Thus, the position of the base $x_oy_oz_o$ frame is given by $\mathbf{r}_o$, and the relative position of each secondary $x_iy_iz_i$ frame is given by $\mathbf{s}_i$. To specify the orientation of these frames, a set of rotational coordinates is defined, after which the components of the vector-dyadic equations of motion can be expressed in a concise matrix form.

4.1 Rotational Coordinates. Several popular sets of rotational coordinates including Euler parameters, Euler-Rodrigues parameters, and the rotation vector rely for their definition on Euler's theorem. This theorem effectively states that the general orientation of a rigid body (or reference frame) can be specified by the angle of rotation θ about a single axis, and the unit vector $\mathbf{u}$ along this axis. A lesser-known set of three parameters is the *conformal rotation vector*, the components of which are also defined in terms of $\mathbf{u}$ and θ:

$$p_i = u_i \tan\frac{\theta}{4};\ i = 1,2,3 \quad (19)$$

where u_i represents the ith component of $\mathbf{u}$ along a base reference frame, and the parameters p_i define the orientation of a secondary frame. The conformal rotation vector was recently introduced by Milenkovic (1982) and employed by Geradin and Cardona (1989) in their work in elastodynamics.

Milenkovic has shown how the rotation matrix $[R]$ can be calculated from the parameters p_i:

$$R_{ij} = \frac{2}{p_o^2 + p_kp_k}\left[p_ip_j + \frac{1}{2}(p_o^2 - p_kp_k)\delta_{ij} - \epsilon_{ijk}p_op_k\right];\ i,j,k = 1,2,3 \quad (20)$$

in which δ_{ij} is the Kronecker delta and ϵ_{ijk} is the alternating tensor. The variable p_o is introduced in order to simplify some of the mathematical expressions involving the conformal rotation vector, and is defined by:

$$2p_o = 1 - p_1^2 - p_2^2 - p_3^2. \quad (21)$$

Using the rotation matrix, one can relate the local and base components of an arbitrary vector $\mathbf{b}$:

$$\{b\} = [R]\{b'\} \text{ or } \{b'\} = [R]^t\{b\} \quad (22)$$

in which $\{b\}$ is the column matrix containing the three base components of $\mathbf{b}$, $\{b'\}$ contains the three local components, and $[R]^t$ is the transpose (i.e., inverse) of the orthonormal rotation matrix. The nine local and base components of an arbitrary dyadic B are related by a similar transformation

$$[B] = [R][B'][R]^t \text{ or } [B'] = [R]^t[B][R]. \quad (23)$$

Milenkovic has also shown how the time derivative of the conformal rotation vector can be expressed in terms of the angular velocity ω of the local frame relative to the base frame:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\dot{p}_i = \frac{1}{2}(p_i p_k + p_o \delta_{ik} + \epsilon_{ijk} p_j)\omega_k'; \; i,j,k = 1,2,3 \tag{24}$$

in which ω_k' represents the kth component of ω in the local frame. Alternatively,

$$\dot{p}_i = \frac{1}{2}(p_i p_k + p_o \delta_{ik} - \epsilon_{ijk} p_j)\omega_k \tag{25}$$

in which ω_k is the kth component of ω in the base frame.

There are several advantages to using the conformal rotation vector in place of an alternate set of parameters. Unlike the set of nine direction cosines or the set of four Euler parameters, only three parameters are required in calculations. A set of Eulerian angles has only three parameters, but the equations corresponding to equation (20) (for the rotation matrix) and equation (24) (for the relationship to the angular velocity) involve transcendental function calls—a distinct numerical disadvantage. Like any other set of three parameters, the conformal rotation vector has a singularity point that can cause numerical difficulties. Unlike other sets, however, there is very little computational expense involved in switching to a complementary set of parameters when this singularity position is approached. Furthermore, the complementary parameters continue to measure orientations relative to the original base frame. If one were to employ Eulerian angles or the Euler-Rodrigues parameters (Roberson, 1985), switching to a new set of angles or parameters would result in subsequent orientations being measured relative to the local frame at the time of the switch. Although this latter approach may prove to be advantageous in some situations (possibly in elastodynamic problems), it would not be so in the current work.

From equation (19), the singularity point for the conformal rotation vector corresponds to a rotation angle θ equal to 2π radians. In numerical calculations, this singularity is avoided by replacing θ with $\theta - 2\pi$ which, according to Euler's theorem, represents the same final body orientation. The components of the complementary set p_i' corresponding to this new angle are calculated using equation (19) and some trigonometric identities:

$$p_i' = u_i \tan\frac{(\theta - 2\pi)}{4} = -u_i\left[\tan\frac{\theta}{4}\right]^{-1}. \tag{26}$$

Using equation (19) again, these complementary components can be expressed directly in terms of the original set of parameters:

$$p_i' = -\frac{p_i}{p_1^2 + p_2^2 + p_3^2} \tag{27}$$

and by substituting this equation back into equation (20), one can verify that $R_{ij}' = R_{ij}$, i.e., the rotation operator is unchanged. In a numerical simulation, the angle θ is not calculated. Instead, equation (27) is used to switch to the complementary set of parameters whenever the inequality

$$p_1^2 + p_2^2 + p_3^2 \le 1 \tag{28}$$

is violated, i.e., when the magnitude of θ exceeds π radians. By enforcing equation (28) during a numerical simulation, the three parameters are guaranteed to remain well behaved.

4.2 Component Form of Equations. Using the transformation equations (22) and (23), the vectors and dyadics appearing in equation (13) can be resolved into components along the primary $x_o y_o z_o$ frame. The resulting three first-order ordinary differential equations can be expressed in symbolic matrix form:

$$[I]\{\dot{\omega}\} = \{M\} \tag{29}$$

in which $\{\dot{\omega}\}$ contains the three components of the angular acceleration α_o of the base body. The column matrix $\{M\}$ is a nonlinear function of the conformal rotation vector and angular velocity of the base body, as well as the physical parameters and relative motions specified for the secondary masses. Finally, $[I]$ contains the nine $x_o y_o z_o$ components of the system's inertia dyadic about G, and is a symmetric positive-definite matrix. As a result, the inverse of $[I]$ can always be calculated in order to solve equation (29) for the angular acceleration of the base vehicle.

Although the conformal rotation vector specifying the relative orientation of each $x_i y_i z_i$ frame is assumed to be a given function of time, the conformal rotation vector defining the absolute orientation of the base $x_o y_o z_o$ frame is unknown. However, it can be expressed as a function of the angular velocity ω_o:

$$\{\dot{p}\} = \frac{1}{2}[H]\{\omega\} \tag{30}$$

where $\{\omega\}$ contains the three $x_o y_o z_o$ components of ω_o. The nine components of $[H]$ are obtained directly from equation (24):

$$[H] = \begin{bmatrix} p_0 + p_1^2 & p_1 p_2 - p_3 & p_1 p_3 + p_2 \\ p_2 p_1 + p_3 & p_0 p + p_2^2 & p_2 p_3 - p_1 \\ p_3 p_1 - p_2 & p_3 p_2 + p_1 & p_0 + p_3^2 \end{bmatrix}. \tag{31}$$

The six differential equations given by equations (29) and (30) represent the six degrees of rotational freedom of the base vehicle.

Resolving equation (5) into components along the inertial XYZ frame, one obtains:

$$\{\dot{v}\} = \frac{\{F\}}{m} - \{\ddot{s}_G\} \tag{32}$$

where $\{\dot{v}\}$ contains the three components of the translational acceleration $\mathbf{a}_o$ of the base mass, and $\{F\}$ contains the components of the resultant external force $\mathbf{F}$. Since the relative acceleration $\ddot{\mathbf{s}}_G$ is a function of α_o given by equations (6) and (7), equation (29) must be solved before the components of $\mathbf{a}_o$ can be obtained from equation (32). The set of governing equations is completed by the simple relationship between the XYZ components of the translational position $\mathbf{r}_o$ and velocity $\mathbf{v}_o$:

$$\{\dot{r}\} = \{v\}. \tag{33}$$

Finally, the $x_i y_i z_i$ components of the internal forces $\mathbf{P}_i$ and moments $\mathbf{N}_i$ can be calculated by applying a similar coordinatization process to the inverse dynamic equations (15) and (17).

5 Computer Program

The matrix equations (29), (30), (32), and (33) represent a set of 12 first-order ordinary differential equations. Since an analytical solution to these highly nonlinear equations is generally out of the question, numerical methods have been employed. A Fortran-77 computer program has been developed to numerically integrate these governing equations using a fourth-order Runge-Kutta algorithm. The program runs under SunOS Unix on a Sun 4 workstation and contains approximately 2000 lines of code organized into three major sections.

In the first section, the parameters necessary to run a numerical simulation are read from the terminal or a data file. In addition to the number of bodies (N) connected to the base mass, this section prompts the user for the motion of each secondary body relative to the base vehicle. The user can specify either the relative position vector and conformal rotation vector of each secondary mass as explicit functions of time, or the relative translational and rotational acceleration of each secondary body. In the latter case, these accelerations are numerically integrated during the simulation to provide the rel-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

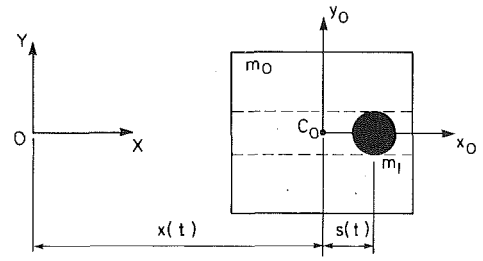


Fig. 3 Translating body carrying an oscillating particle

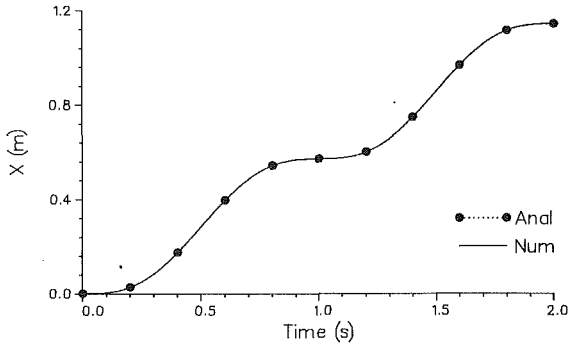


Fig. 4 Position of base mass

ative position and orientation of each secondary mass at each time step.

This section also prompts for the twelve initial conditions of the base vehicle, i.e., the position and velocity at the start of the simulation. For each secondary mass in the system the inertial properties m_i and I_i are input, as well as the external loads $\mathbf{F}_i$ and $\mathbf{M}_i$ and the location of the contact point $\mathbf{q}_i$ with the base vehicle. Finally, the user is prompted for the initial and final simulation times, the integration time step, and the number of time steps between each output of numerical results.

The second section contains the Runge-Kutta algorithm which numerically integrates the first time derivatives contained in the left-hand sides of the governing differential equations. At each time step of a simulation, equation (29) is solved for $\{\dot{\omega}\}$ using a Cholesky algorithm to take advantage of the symmetric positive-definite nature of $[I]$. Equations (30), (32), and (33) are then solved for the remaining first derivatives. These derivatives are numerically integrated to get new values for the variables at the start of the next step. At each output interval, the internal forces and moments are calculated and the numerical results of interest are written to output arrays. When the numerical simulation has completed, the third section writes these output arrays to a file on disk. Work is currently in progress to allow the user to produce X-Y plots of the numerical results or to generate computer graphic images of solid models representing the rigid bodies in the system.

This computer code can be classified as a special-purpose multibody program capable of analyzing a system with a "dandelion" topology. The stem of the dandelion graph represents the connection between the base vehicle and the inertial frame of reference, while each dandelion seed represents the connection between a secondary mass and the base vehicle. The only possible topological difference between two modeled systems is the number (N) of secondary masses. Unlike a *general-purpose* multibody formulation (e.g., Roberson and Schwertassek, 1988; Li and Andrews, 1986) therefore, the equations of motion presented in this paper were derived without the application of linear graph theory. Mathematically, the only effect of a different system topology on these equations is a change in the upper limit N on the various summations. Before this discussion is carried too far, however, it should be re-emphasized that the main goal of this paper is to illustrate the dynamic effects of known configuration changes, and not to present another multibody formulation.

6 Illustrative Examples

6.1 One-Dimensional Problem. As shown in Fig. 3, the physical system consists of a vehicle m_o containing an oscillating particle m_1, both of which are translating along the inertial X-axis. The relative position of the particle is assumed to be a simple harmonic function of time with an amplitude of one meter:

$$s(t) = \sin 2\pi t. \tag{34}$$

The mass of the particle is one kilogram and the mass of the vehicle is ten kilograms. Assuming that there are no external loads on the system ($F_o = F_1 = 0$) and that the base vehicle starts from rest, the effect of the oscillating particle on the motion of the base mass will be determined. This one-dimensional model might be used as a simple representation of a transport truck carrying a shifting load.

Since there are no external forces on the system, an analytical solution can be effected by applying the law of conservation of linear momentum. At any time t, the linear momentum L is given by:

$$L = m_o v + m_1(v + \dot{s}) \tag{35}$$

and initially, its value is:

$$L(0) = m_1 \dot{s}(0) = 2\pi m_1 \tag{36}$$

Equating these two expressions for L, and solving for the velocity of the base mass,

$$v(t) = \frac{2\pi m_1}{m_0 + m_1}\,[1 - \cos 2\pi t]. \tag{37}$$

which can be integrated to obtain the position of the vehicle:

$$x(t) = \frac{2\pi m_1}{m_o + m_1}\left[t - \frac{\sin 2\pi t}{2\pi}\right]. \tag{38}$$

This analytical solution is plotted in Fig. 4, along with the results from the numerical simulation of this system. Using a timestep of 0.01 seconds, the numerical results agree with the analytical solution to at least eight significant figures throughout the two seconds of simulation.

The internal force P driving the oscillating particle can also be solved for analytically. Applying Newton's second law:

$$P = m_1 a_1 = m_1(a_o + \ddot{s}) \tag{39}$$

where a_o is the acceleration of the base mass:

$$a_o = \dot{v} = \frac{4\pi^2 m_1}{m_o + m_1} \sin 2\pi t. \tag{40}$$

Thus, the actuator force responsible for the relative motion of the particle is given by:

$$P = -\frac{4\pi^2 m_0 m_1}{m_o + m_1} \sin 2\pi t. \tag{41}$$

This analytical solution is plotted in Fig. 5, along with the numerical results from the encoded inverse dynamic formulation.

6.2 Three-Dimensional Problem. In this example, the physical system consists of a spinning satellite with four folding solar panels. The satellite base has been modeled as an axisymmetric rigid body m_o, and each panel has been represented as a thin plate m_i that is connected to the base body by a revolute joint. Figure 6 shows the initial configuration of the system in which the base $x_o y_o z_o$ frame is coincident with the inertial XYZ frame, and each solar panel is fully deployed. The initial velocity of C_o is assumed to be zero, and the satellite

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

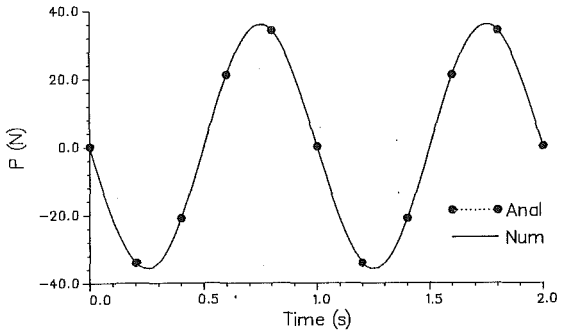


Fig. 5 Internal force on particle

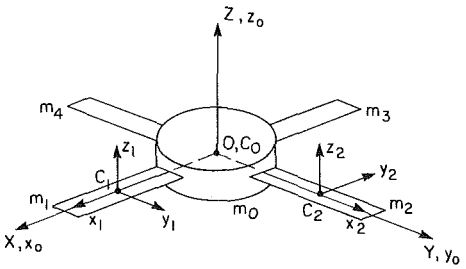


Fig. 6 Spinning satellite with folding panels

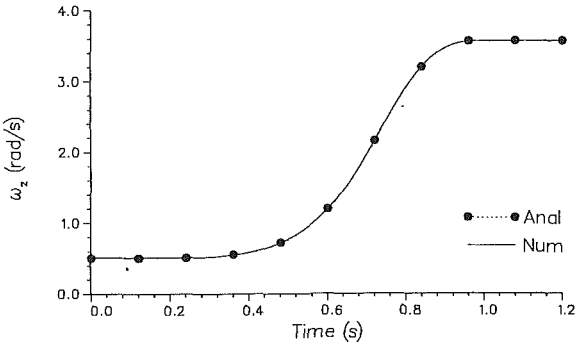


Fig. 7 Spin rate of base body

is spinning about the vertical Z-axis at the start of the simulation. The purpose of this problem is to determine the subsequent rotational motion of the base satellite if the four panels are rotated through 90 degrees into an upright position. Hibbeler (1986) has investigated this problem previously, and solved analytically for the final spin rate for the case in which the four panels fold up simultaneously. Using the formulation presented in this paper, it is possible to solve for the angular velocity of the base body at any time during the motion, even for the case in which the panels do not fold up simultaneously.

The relative orientation of each panel can be defined by the angle θ_i between the x_i-axis and the $x_o y_o$ plane, where $0 \leq \theta_i \leq 90$ deg. The relative rotational motion of each panel can then be specified in terms of the given angle θ_i:

$$\omega_{i/o} = -\dot{\theta}_i \mathbf{j}_i, \quad \mathring{\omega}_{i/o} = -\ddot{\theta}_i \mathbf{j}_i \tag{42}$$

where $\mathbf{j}_i$ is the unit vector along the secondary y_i-axis. In $x_o y_o z_o$ components, these become:

$$\omega_{1/o} = -\dot{\theta}_1 \mathbf{j}_o, \ \omega_{2/o} = \dot{\theta}_2 \mathbf{i}_o, \ \omega_{3/o} = \dot{\theta}_3 \mathbf{j}_o, \ \omega_{4/o} = -\dot{\theta}_4 \mathbf{i}_o$$
$$\mathring{\omega}_{1/o} = -\ddot{\theta}_1 \mathbf{j}_o, \ \mathring{\omega}_{2/o} = \ddot{\theta}_2 \mathbf{i}_o, \ \mathring{\omega}_{3/o} = \ddot{\theta}_3 \mathbf{j}_o, \ \mathring{\omega}_{4/o} = -\ddot{\theta}_4 \mathbf{i}_o \tag{43}$$

in which $\mathbf{i}_o$, $\mathbf{j}_o$, and $\mathbf{k}_o$ are the unit vectors along the x_o, y_o, and z_o-axes, respectively. The relative translational position of C_i can also be expressed in terms of θ_i:

$$\mathbf{s}_1 = (a + b\cos\theta_1)\mathbf{i}_o + b\sin\theta_1 \mathbf{k}_o, \ \mathbf{s}_2 = (a + b\cos\theta_2)\mathbf{j}_o + b\sin\theta_2 \mathbf{k}_o$$
$$\mathbf{s}_3 = -(a + b\cos\theta_3)\mathbf{i}_o + b\sin\theta_3 \mathbf{k}_o,$$
$$\mathbf{s}_4 = -(a + b\cos\theta_4)\mathbf{j}_o + b\sin\theta_4 \mathbf{k}_o \tag{44}$$

where a is the radius of the cylindrical base, and b is one half the length of each uniform panel. The relative velocity $\mathring{\mathbf{s}}_i$ and acceleration $\mathring{\mathring{\mathbf{s}}}_i$ are obtained by differentiating the scalar components of equations (44)—the unit vectors $\mathbf{i}_o$, $\mathbf{j}_o$, and $\mathbf{k}_o$ being constants under this derivative operator.

The following parameter values have been assumed for the analysis:

$$m_o = 125 \text{ kg}, \ m_i = 20 \text{ kg}, \ a = 0.2\text{m}, \ b = 0.375 \text{ m}$$
$$I_{xx_o} = I_{yy_o} = I_{zz_o} = 0.94 \text{ kg}\cdot\text{m}^2 \tag{45}$$
$$I_{xx_i} = 0.0667 \text{ kg}\cdot\text{m}^2, \ I_{yy_i} = 0.9375 \text{ kg}\cdot\text{m}^2, \ I_{zz_i} = 1.0042 \text{ kg}\cdot\text{m}^2$$

in which the parameters for each panel m_i correspond to a thin plate measuring 0.2 m by 0.75 m. These values have been taken directly from Hibbeler's problem with the exception of I_{xx_o} and I_{yy_o}. These two parameters were not specified by Hibbeler because they are not required if the panels fold up simultaneously. The satellite is assumed to be far enough from any planetary mass that the external forces and moments are zero. At the start of the simulation, the satellite is assumed to be spinning at 0.5 radians per second about the z_o-axis.

For the case in which the panels do fold up simultaneously, the following function is given for the orientation of each panel:

$$\theta_1 = \theta_2 = \theta_3 = \theta_4 = \frac{\pi}{2} t - \frac{\sin 2\pi t}{4} \tag{46}$$

so that θ_i is initially zero and rotates through $\pi/2$ radians in one second. After this second, the panels are assumed to lock in the upright position. By differentiating equation (46), one can show that this specified angular motion provides a smooth start and finish, i.e., the relative acceleration is zero at the beginning and end of this second.

Since the z_o and Z-axis remain parallel as the panels fold up symmetrically, the angular momentum of the system about its center of mass G can be expressed as a function of ω_z, the component of ω_o about the spin axis:

$$\mathbf{h}_G = [80\omega_z(0.2 + 0.375\cos\theta_i)^2$$
$$+ 4\omega_z(I_{xx_i}\sin^2\theta_i + I_{zz_i}\cos^2\theta_i) + 0.94\omega_z]\mathbf{k}_o \tag{47}$$

which takes the initial value:

$$\mathbf{h}_G(0) = 15.703\mathbf{k}_o. \tag{48}$$

Since there is no external moment about the center of mass, conservation of this angular momentum can be applied in order to solve for ω_z:

$$\omega_z = \frac{15.703}{80(0.2 + 0.375\cos\theta_i)^2 + 4(I_{xx_i}\sin^2\theta_i + I_{zz_i}\cos^2\theta_i) + 0.94}. \tag{49}$$

This analytical solution is plotted in Fig. 7, along with the results from a numerical simulation. The two curves are practically indistinguishable, and show that the spin of the satellite increases more than sevenfold as the panels fold up. Once the panels have locked into position, the spin rate remains constant at 3.56 radians per second, the value predicted by Hibbeler. The effect of the panel motions is to alter the system's configuration so as to reduce the moment of inertia about the spin axis. Since the angular momentum about this axis is conserved, the rate of spin increases. This effect is analogous to that employed by figure skaters to increase their spin rate during a maneuver by crossing their arms against their chest.

If the panels do not fold up simultaneously, the local z_o-axis will not remain parallel to the inertial Z-axis because the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

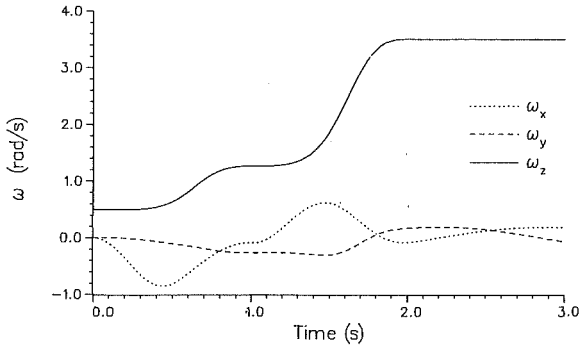


Fig. 8 Angular velocity of base body

satellite will "wobble." Conservation of angular momentum can still be applied, but the result is three nonlinear equations which cannot (in general) be solved explicitly for the three unknown components ω_x, ω_y, and ω_z of the angular velocity of the base body. A numerical solution can be effected, however, using the encoded formulation presented in this paper. It will be assumed that the first three panels fold up in the first second of simulation, as before. The fourth panel, however, is specified to fold up only after the first three have locked into position:

$$\theta_4 = \frac{\pi}{2}(t-1) - \frac{\sin 2\pi t}{4}, \quad 1 \le t \le 2. \tag{50}$$

The same system parameters and initial conditions have been employed in a numerical simulation, the results of which are plotted in Fig. 8.

The component of angular velocity about the z_o-axis, ω_z, increases over the first second as the first three panels fold up, then increases even further as the fourth panel is retracted. After all panels have locked into place, the motion of the system corresponds to the coning motion of a torque-free axisymmetric body. Thus, the component ω_z about the symmetry axis maintains a constant value, while the components ω_x and ω_y assume sinusoidal functions of time. The final value of ω_z is 3.514 radians per second, a value slightly less than the previous final spin rate. The explanation for this is as follows. The component of the angular momentum of the system about the inertial Z-axis is a constant. For the given parameters, the constant angle between this Z-axis and the z_o-axis is only about 9.57 deg once the panels have folded up. Thus, the majority of the angular momentum about the Z-axis is due to the constant component ω_z, but a small portion is also provided by the nonzero components ω_x and ω_y.

An analytical solution for the rotational motion of this system is only possible once the panels have all been retracted. Therefore, verification of the numerical results for the first two seconds of simulation was not a straightforward procedure. One check was provided by calculating the angular momentum of the system $\mathbf{h}_G$ at each output interval using equation (12). The fact that this angular momentum remained constant to nine significant figures after 3,000 integration steps provided some assurance that the results were valid. A more convincing check was provided by showing that the results satisfied the principle of work and energy. To do so, the internal torques responsible for folding up the panels were calculated using the encoded inverse dynamic analysis. The work done on the system by these four different torques was then calculated by numerically integrating the torques with respect to the relative angular position of the corresponding panel. The kinetic energy of the system was calculated initially and after two seconds using:

$$KE = \frac{1}{2}\sum_{i=0}^{N}[m_i\mathbf{v}_i\cdot\mathbf{v}_i + \boldsymbol{\omega}_i\cdot\boldsymbol{I}_i\cdot\boldsymbol{\omega}_i]. \tag{51}$$

By showing that the change in kinetic energy was equal to the work done by the internal torques, the work-energy principle was found to be satisfied.

7 Conclusion

A vector-dyadic formulation of the dynamic equations of motion for a system with time-varying inertial properties has been presented. The system has been represented by the multibody model proposed by Roberson (1958) and subsequently employed by Grubin (1962). This model is general enough to represent a wide variety of physical systems while remaining mathematically tractable. The governing equations of motion account for the time-varying nature of the system's inertia dyadic, as well as the changing location of the center of mass within the system.

This formulation has been encoded into a computer program for the purpose of numerically simulating the dynamics of these nonlinear systems. Essentially, this computer code can be classified as a special-purpose multibody program. Rotations have been described using the conformal rotation vector—a set of three parameters that give rise to algebraic expressions for the rotation operator and the angular velocity vector. Finally, two example simulations were presented to clearly illustrate the dominating effect that configuration changes may have on a system, and the ability of the encoded formulation to model these effects.

References

DeBra, D. B., 1963, "Vectors and Dyadics: The Natural Tools of Space-Vehicle Mechanics," *Advances in the Astronautical Sciences*, Vol. 9, pp. 77–107.

Ge, Z.-M., 1984, "The Equations of Motion of Nonlinear Nonholonomic Variable Mass System with Applications," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 435–437.

Geradin, M., and Cardona, A., 1989, "Kinematics and Dynamics of Rigid and Flexible Mechanisms Using Finite Elements and Quaternion Algebra," *Computational Mechanics*, Vol. 4, pp. 115–136.

Grubin, C., 1962, "Dynamics of a Vehicle Containing Moving Parts," ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, pp. 486–488.

Grubin, C., 1963, "Mechanics of Variable Mass Systems," *Journal of the Franklin Institute*, Vol. 276, pp. 305–312.

Hibbeler, R. C., 1986, *Engineering Mechanics: Statics and Dynamics*, 4th ed., Macmillan, New York.

Kapoulitsas, G., 1987, "The Mass-centre Motion of a Continuously Variable System of Particles—Part 2," *Ingenieur-Archiv.*, Vol. 57, pp. 91–98.

Keat, J. E., and Turner, J. D., 1987, "Equations of Motion of Systems of Variable-Mass Bodies for Space Structure Deployment Simulation," *Proceedings, The Sixth VPI&SU/AIAA Symposium on Dynamics and Control of Large Structures*, Blacksburg, Va.

Li, T. W., and Andrews, G. C., 1986, "Application of the Vector-Network Method to Constrained Mechanical Systems," ASME *Journal of Mechanisms, Transmission, and Automation in Design*, Vol. 29, pp. 471–480.

McPhee, J. J., 1989, "Dynamics of Systems with Time-Varying Inertial Properties," *Proceedings, The Twelfth Canadian Congress of Applied Mechanics*, Ottawa, Ontario, pp. 392–393.

Meirovitch, L., 1970, *Methods of Analytical Dynamics*, McGraw-Hill, New York.

Milenkovic, V., 1982, "Coordinates Suitable for Angular Motion Synthesis in Robots," *Proc. Robots VI*, (RI of SME), pp. 407–420.

Roberson, R. E., 1958, "Torques on a Satellite Vehicle From Internal Moving Parts," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 196–200.

Roberson, R. E., 1985, "On the Practical Use of Euler-Rodrigues Parameters in Multibody System Dynamic Simulation," *Ingenieur-Archiv*, Vol. 55, pp. 114–123.

Roberson, R. E., and Schwertassek, R., 1988, *Dynamics of Multibody Systems*, Springer-Verlag, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

W. Jiang
Assistant Professor,
Mechanical Engineering Department.

W. K. Jones
Associate Professor,
Mechanical Engineering Department.

T. L. Wang
Associate Professor,
Civil Engineering Department.

K. H. Wu
Associate Professor,
Mechanical Engineering Department.

Florida International University,
The State University of Florida at Miami,
Miami, FL 33199

Free Vibration of Helical Springs

This paper presents a theoretical investigation of the coupled extensional-torsional vibration of helical springs. The study shows that two types of periodic wave will propagate through the spring, one characterizing the extensional-compressive deformation and the other one, the torsional deformation. The shapes of the individual waves are simple, but the oscillation of the spring is complex due to the interaction and superposition of the component waves.

I Introduction

In addition to many practical engineering problems involving the dynamics of springs, many textbooks use the spring, together with the damper, to illustrate the theory of mechanical vibrations and the relevant applied mathematics. In both practice and theory, a spring is usually considered as a linear extension-compression element, the deformation of which is proportional to the force applied on it, with a constant of proportionality called the spring constant. In fact, the spring is a nonlinear mechanical element, and the extensional-compressive deformation of the spring is always accompanied by the torsional deformation, and vice versa (Jiang et al., 1989). Such coupled behavior changes the dynamic response of the spring. The study of this phenomenon is important, not only because of the many applications of springs in common engineering practice, but also because of the requirement of maximum reliability in certain situations, such as in the use of springs as implantable cardiac leads in biomechanics. Additionally, the stress and strain analysis of wire ropes can also be based on the spring theory (Phillips and Costello, 1977), and the study of wire rope oscillations is in increasing demand, for example, in ocean salvage and construction operations and in deep-water moorings.

The derivation of the governing differential equations for helical spring dynamics can be dated back as early as 1890, when Michell obtained three equations of motion by using the first form of Lagrange's fundamental equation. In the derivation, Michell considered displacements in three directions, but only one independent rotation. All the moments of inertia of the typical element and the deformation produced by all forces were ignored. He showed the existence of two types of wave, producing axial or radial movement. These two waves were independent.

Love (1899) obtained six equations of motion based on the same assumptions as Michell's. His equations were later modified by Yoshimura and Murata (1952) to include the torsional inertia, and then by Wittrick (1966) to include the rotary inertia and Timoshenko shear deformation effects. The governing equations thus derived, however, are too difficult to permit analytical solutions. Consequently, various numerical methods, together with a variety of simplifications, were introduced to attack such problems (e.g., Mottershead, 1980; Pearson and Wittrick, 1986). Interested readers can refer to an extensive literature review by Pearson (1988).

Some "simple" theories were also developed along with the above "complex" theories. The simplest one treated the spring as a homogeneous rod having modified elastic and mass properties (e.g., Dick, 1957; Gross, 1959; Curran, 1970). The coupling effects relating extension, torsion, and bending were ignored. Phillips and Costello (1972) extended the simple theory to include the extensional-torsional coupling effect for springs under the action of axial forces and moments, and a set of nonlinear equations was obtained. Costello (1975) further extended the investigation in the radial expansion due to axial impact and the linear and nonlinear theories were com-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 6, 1989; final revision, Feb. 20, 1990.
Paper No. 91-APM-4.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

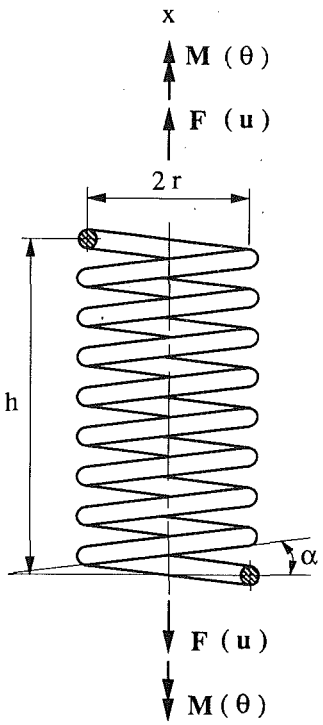


Fig. 1 Helical spring subjected to end forces and torques

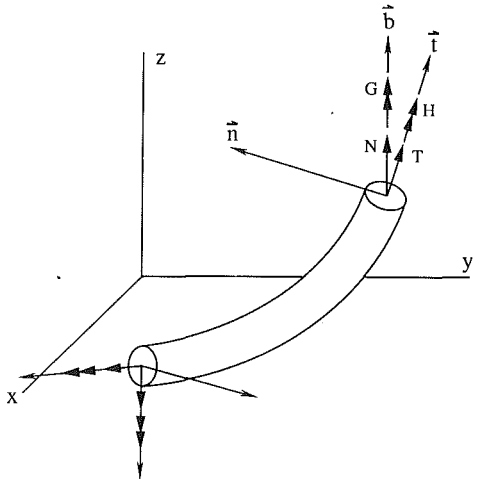


Fig. 2 Resultants that act on a wire element

pared in a paper by Sinha and Costello (1978). Similar linear equations of motion were used by Samras et al. (1974) in the study of wire rope oscillations and by Phillips and Costello (1977) in the investigation of the axial impact of twisted wire cables.

The background review shows that while complex theories take more factors into consideration to deliberate the governing equations, the complexity of these equations makes it practically impossible to obtain an exact analytical solution, and the numerical methods used instead permit only the calculations of natural frequencies and normal modes. The interesting coupling effects of how the extension, torsion, and bending deformations interact and the important problem of how resonance occurs remain unsolved. To gain a deep insight into such problems, analytical solutions are needed, and analytical solutions can only be expected for some simple theories.

This paper presents an analytical investigation of the coupled extensional-torsional dynamic behavior of the helical springs, based on the simple theory introduced by Phillips and Costello (1972) and the force-strain relationships we derived in a previous paper (Jiang et al., 1989). As an example, the free vibration of the spring is worked out, and a closed-form solution is found. The solution shows that two types of periodic wave will propagate through the spring, one characterizing the extensional-compressive deformation and the other one the torsional deformation. The shapes of the individual waves are simple, but the oscillation of the spring is complex and not periodic due to the interaction and superposition of the two different waves. The study further shows that the interaction between these two types of wave depends on the helical angle of the spring. When the helical angle approaches 90 deg, the interaction becomes weaker and weaker and finally disappears, and the spring behaves like a rod. As a by-product, this paper obtains the closed-form solution to the longitudinal free vibration of a rod. Our solution shows that the longitudinal wave propagating through the rod is of a very simple trapezoidal shape from which very detailed information about the process of wave propagation through the rod can be derived. The longitudinal vibration of a rod is a classic problem that appears in almost all the vibration and applied mathematics textbooks. However, it appears that no one has obtained such a closed-form solution before, and based on these results, we believe that this paper also presents some interesting results in the theory of vibrations.

II Equations of Motion

The force-strain relationships and the equations of motion of helical springs have been derived in a previous paper (Jiang et al., 1989). This paper concerns the coupled extensional-torsional free vibration of helical springs having constant helical angle and helix radius under small deformation.

According to Love (1927), a helical spring, if there are no distributed loads applied on it, can be deformed into the shape of another helix under axial force F and twist moment M, which form a wrench at each end, of which the axis is the axis of the helix (Fig. 1). In such a case, the nonzero internal resultant forces and moments are constant along the spring wire center line, and they are the tensile force T, the binormal shear force N, the binormal bending moment G, and the torsional couple H (Fig. 2).

The linear force-strain relationships have been found as

$$\begin{cases} F = k_1\epsilon + k_2\phi \\ M = k_3\epsilon + k_4\phi \end{cases} \tag{1}$$

where ϵ is the axial strain and ϕ is the rotational strain of the spring, and

$$\begin{cases} k_1 = \dfrac{\pi E R^4}{r^2\Delta}\,[1 + k^2(1+\nu)] \\ k_2 = -\dfrac{\pi\nu E R^4 k}{r\Delta} \\ k_3 = -\dfrac{\pi E R^4 k}{4r\Delta}\left(4\nu + \dfrac{R^2}{r^2}\right) \\ k_4 = \dfrac{\pi E R^4}{4\Delta}\left[4(1+\nu+k^2) + \dfrac{k^2R^2}{r^2}\right] \end{cases} \tag{2}$$

are the stiffness constants. In equation (2), R is the wire radius, r is the helix radius, k is the tangent of the helical angle α, E is Young's modulus, ν is the Poisson's ratio, and

$$\Delta = \frac{1}{k\sqrt{1+k^2}}\left\{4(1+\nu)(1+k^2)^2 + \frac{k^2R^2}{r^2}\,[(1-\nu)+k^2(1+\nu)]\right\} \tag{3}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Based on the linear force-strain relationships, equation (1), the equations of motion for the small deformations of the spring can be found as:

$$\begin{cases} k_1 \dfrac{\partial^2 u}{\partial x^2} + k_2 \dfrac{\partial^2 \theta}{\partial x^2} = \gamma \dfrac{\partial^2 u}{\partial t^2} \\ k_3 \dfrac{\partial^2 u}{\partial x^2} + k_4 \dfrac{\partial^2 \theta}{\partial x^2} = \mu \dfrac{\partial^2 \theta}{\partial t^2} \end{cases} \tag{4}$$

where $u(x, t)$ and $\theta(x, t)$ are the axial and rotational displacement at time t (Fig. 1), and γ is the mass and μ is the mass moment of inertia about the axis per unit length of the spring in the unloaded configuration.

It is known that in order for a spring to have a positive-definite strain energy function, or, in the dynamic analysis, in order for a conservative system to associate with a self-adjoint problem, the stiffnesses of the spring should satisfy the relations:

$$k_2 = k_3 \tag{5}$$

and

$$k_1 k_4 - k_2 k_3 > 0. \tag{6}$$

Now the stiffnesses, equation (2), do satisfy the second requirement, equation (6), but not the first one, equation (5). Some comments are as follows:

The linear force-strain relationships, equation (1), are the result of a linearization of the nonlinear, force-strain relationships obtained in a previous paper (Jiang et al., 1989). The linearization has been carefully checked and the stiffness expressions, equation (2), prove to be correct. It should be noted that similar force-strain relationships can also be found for cable and wire rope structures. The stiffnesses k_2 and k_3 of a wire rope obtained by Samras et al. (1974) by means of measurement are different. Also, Costello (1983) obtained different values of these stiffnesses based on their theory. In other words, this kind of conflict with the self-adjoint requirement is not uncommon in the spring and wire rope structures. On the other hand, the difference between k_2 and k_3 in most cases is found to be very small. In fact, the term $\dfrac{R^2}{r^2}$ in the expression of k_3 is generally negligible in comparison with the term 4ν, and when that term is neglected, the stiffnesses k_2 and k_3 do become equal. Since there is no reason to artificially make them exactly the same, we will use these stiffness constants in this paper with the above comments in mind.

III Solution by Laplace Transforms

The linearized equations of motion, equation (4), can be solved using, for example, the method of Laplace transform, which is generally valid for arbitrary initial and boundary conditions.

With the Laplace transforms of $u(x, t)$ and $\theta(x, t)$ denoted by $\bar{u}(x, s)$ and $\bar{\theta}(x, s)$, respectively, the equations of motion, equation (4), can be transformed to give

$$\begin{cases} k_1 \dfrac{\partial^2 \bar{u}}{\partial x^2} + k_2 \dfrac{\partial^2 \bar{\theta}}{\partial x^2} - \gamma s^2 \bar{u} = \gamma u_0(x) \\ k_3 \dfrac{\partial^2 \bar{u}}{\partial x^2} + k_4 \dfrac{\partial^2 \bar{\theta}}{\partial x^2} - \mu s^2 \bar{\theta} = \mu \theta_0(x) \end{cases} \tag{7}$$

where

$$\begin{cases} u_0(x) = -su(x,0) - \dfrac{\partial}{\partial t} u(x,0) \\ \theta_0(x) = -s\theta(x,0) - \dfrac{\partial}{\partial t} \theta(x,0) \end{cases} \tag{8}$$

are known functions given by the initial conditions of the problem.

Assuming that the solution is of the form

$$\begin{cases} \bar{u}(x,s) = f(s) e^{\frac{\omega x s}{h}} + \bar{u}^*(x) \\ \bar{\theta}(x,s) = g(s) e^{\frac{\omega x s}{h}} + \bar{\theta}^*(x) \end{cases} \tag{9}$$

where $(\bar{u}^*, \bar{\theta}^*)$ is a particular solution, and (f, g) is the solution to the homogeneous equations:

$$\begin{cases} (k_1 \omega^2 - \gamma h^2) f(s) + k_2 \omega^2 g(s) = 0 \\ k_3 \omega^2 f(s) + (k_4 \omega^2 - \mu h^2) g(s) = 0, \end{cases} \tag{10}$$

one finds that the parameter ω should satisfy the following determinantal equation

$$(k_1 k_4 - k_2 k_3)\omega^4 - (\mu k_1 + \gamma k_4) h^2 \omega^2 + \gamma \mu h^4 = 0 \tag{11}$$

if the functions $f(s)$ and $g(s)$ are not identically zero. Equation (11) gives roots

$$\begin{cases} \omega_1^2 = \dfrac{h^2[(\mu k_1 + \gamma k_4) - \sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma \mu k_2 k_3}]}{2(k_1 k_4 - k_2 k_3)} \\ \omega_2^2 = \dfrac{h^2[(\mu k_1 + \gamma k_4) + \sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma \mu k_2 k_3}]}{2(k_1 k_4 - k_2 k_3)} \end{cases} \tag{12}$$

and equation (10) gives the ratios of the functions $f(s)$ to $g(s)$. The general solution, equation (9), thus becomes

$$\begin{cases} \bar{u}(x,s) = f_1(s) e^{-\frac{\omega_1 x s}{h}} + f_2(s) e^{\frac{\omega_1 x s}{h}} \\ \qquad + f_3(s) e^{-\frac{\omega_2 x s}{h}} + f_4(s) e^{\frac{\omega_2 x s}{h}} + \bar{u}^*(x) \\ \bar{\theta}(x,s) = \alpha_1 \left[f_1(s) e^{-\frac{\omega_1 x s}{h}} + f_2(s) e^{\frac{\omega_1 x s}{h}} \right] \\ \qquad + \alpha_2 \left[f_3(s) e^{-\frac{\omega_2 x s}{h}} + f_4(s) e^{\frac{\omega_2 x s}{h}} \right] + \bar{\theta}^*(x) \end{cases} \tag{13}$$

where

$$\begin{cases} \alpha_1 = -\dfrac{1}{2\mu k_2} \left[(\mu k_1 - \gamma k_4) - \sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma \mu k_2 k_3} \right] \\ \alpha_2 = -\dfrac{1}{2\mu k_2} \left[(\mu k_1 - \gamma k_4) + \sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma \mu k_2 k_3} \right]. \end{cases} \tag{14}$$

The functions $f_i(s)$ in the general solution, equation (13), will be determined from the boundary conditions, and finally, an inverse Laplace transform gives the solution to the equations of motion.

IV Free Vibration Problems

As an example, consider a spring, which is fixed at one end, $x = 0$, and free at the other end, $x = h$. At first, the spring is subjected to certain initial displacements u_0 and θ_0 at the free end. Then, at time $t = 0$, the initial displacements are suddenly removed with zero initial velocities, and the free vibration of the spring is analyzed.

The initial and boundary conditions for such a problem are as follows:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\begin{cases} u(x,0) = \dfrac{u_0}{h}\, x \\ \theta(x,0) = \dfrac{\theta_0}{h}\, x \end{cases} \tag{15}$$

$$\begin{cases} \dfrac{\partial}{\partial t}\, u(x,0) = 0 \\ \dfrac{\partial}{\partial t}\, \theta(x,0) = 0 \end{cases} \tag{16}$$

$$\begin{cases} u(0,t) = 0 \\ \theta(0,t) = 0 \end{cases} \tag{17}$$

$$\begin{cases} \dfrac{\partial}{\partial x}\, u(h,t) = 0 \\ \dfrac{\partial}{\partial x}\, \theta(h,t) = 0. \end{cases} \tag{18}$$

With these conditions, the general solution, equation (13) yields:

$$\begin{cases} \bar{u}(x,s) = u_0\left[\dfrac{x}{hs} - \displaystyle\sum_{i=1}^{2} \frac{a_i}{\omega_i s^2 \cosh \omega_i s} \sinh \frac{\omega_i x s}{h}\right] \\ \qquad\qquad - \theta_0 \displaystyle\sum_{i=1}^{2} \frac{c_i}{\omega_i s^2 \cosh \omega_i s} \sinh \frac{\omega_i x s}{h} \\ \bar{\theta}(x,s) = -u_0 \displaystyle\sum_{i=1}^{2} \frac{b_i}{w_i s^2 \cosh \omega_i s} \sinh \frac{\omega_i x s}{h} \\ \qquad\qquad + \theta_0\left[\dfrac{x}{hs} - \displaystyle\sum_{i=1}^{2} \frac{d_i}{\omega_i s^2 \cosh \omega_i s} \sinh \frac{\omega_i x s}{h}\right] \end{cases} \tag{19}$$

where

$$\begin{cases} a_1 = \dfrac{1}{2}\left[1 + \dfrac{\mu k_1 - \gamma k_4}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}}\right] \\ a_2 = \dfrac{1}{2}\left[1 - \dfrac{\mu k_1 - \gamma k_4}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}}\right] \end{cases} \tag{20}$$

$$\begin{cases} b_1 = \dfrac{\gamma k_3}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}} \\ b_2 = -\dfrac{\gamma k_3}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}} \end{cases} \tag{21}$$

$$\begin{cases} c_1 = \dfrac{\mu k_2}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}} \\ c_2 = -\dfrac{\mu k_2}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}} \end{cases} \tag{22}$$

$$\begin{cases} d_1 = \dfrac{1}{2}\left[1 - \dfrac{\mu k_1 - \gamma k_4}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}}\right] \\ d_2 = \dfrac{1}{2}\left[1 + \dfrac{\mu k_1 - \gamma k_4}{\sqrt{(\mu k_1 - \gamma k_4)^2 + 4\gamma\mu k_2 k_3}}\right] \end{cases} \tag{23}$$

Then, an inverse Laplace transform gives the solution to the equations of motion, equation (4), the free vibration of a spring under the initial disturbance, equations (15) and (16):

$$\begin{cases} u(x,t) = \displaystyle\sum_{i=1}^{2} (u_0 a_i + \theta_0 c_i)\, w_i(x,t) \\ \theta(x,t) = \displaystyle\sum_{i=1}^{2} (u_0 b_i + \theta_0 d_i)\, w_i(x,t) \end{cases} \tag{24}$$

where

$$w_i(x,t) = \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{(-)^n}{(2n+1)^2} \sin \frac{(2n+1)\pi x}{2h} \cos \frac{(2n+1)\pi t}{2\omega_i}. \tag{25}$$

Note that if the spring is only subjected to an axial initial displacement u_0 at the end and can rotate freely, the spring will have an initial rotation $\theta_0 = -\dfrac{k_3}{k_4} u_0$ due to the coupled behavior. The solution to this special case has been derived in the previous paper (Jiang et al., 1989).

Similarly, if the spring is initially subjected to a force F_0 and a torque M_0 at the end, $x = h$, and then is released without initial velocity, the solution

$$\begin{cases} u(x,t) = \displaystyle\sum_{i=1}^{2} (F_0 e_i + M_0 g_i)\, w_i(x,t) \\ \theta(x,t) = \displaystyle\sum_{i=1}^{2} (F_0 f_i + M_0 h_i)\, w_i(x,t) \end{cases} \tag{26}$$

where

$$\begin{cases} e_1 = \dfrac{hs_1}{2}\left[1 - \dfrac{s_1(\gamma s_1 - \mu s_4) + 2\mu s_2 s_3}{s_1\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \\ e_2 = \dfrac{hs_1}{2}\left[1 + \dfrac{s_1(\gamma s_1 - \mu s_4) + 2\mu s_2 s_3}{s_1\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \end{cases} \tag{27}$$

$$\begin{cases} f_1 = \dfrac{hs_3}{2}\left[1 - \dfrac{\gamma s_1 + \mu s_4}{\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \\ f_2 = \dfrac{hs_3}{2}\left[1 + \dfrac{\gamma s_1 + \mu s_4}{\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \end{cases} \tag{28}$$

$$\begin{cases} g_1 = \dfrac{hs_2}{2}\left[1 - \dfrac{\gamma s_1 + \mu s_4}{\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \\ g_2 = \dfrac{hs_2}{2}\left[1 + \dfrac{\gamma s_1 + \mu s_4}{\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \end{cases} \tag{29}$$

$$\begin{cases} h_1 = \dfrac{hs_4}{2}\left[1 + \dfrac{s_4(\gamma s_1 - \mu s_4) - 2\gamma s_2 s_3}{s_4\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \\ h_2 = \dfrac{hs_4}{2}\left[1 - \dfrac{s_4(\gamma s_1 - \mu s_4) - 2\gamma s_2 s_3}{s_4\sqrt{(\gamma s_1 - \mu s_4)^2 + 4\gamma\mu s_2 s_3}}\right] \end{cases} \tag{30}$$

In equations (27)–(30), s_i, i = 1–4, are slenderness constants of the spring:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\begin{cases} s_1 = \dfrac{1}{\pi ER^2 k\sqrt{1+k^2}}\left[k^2 + \dfrac{4(1+\nu+k^2)r^2}{R^2}\right] \\ s_2 = \dfrac{4\nu r}{\pi ER^4\sqrt{1+k^2}} \\ s_3 = \dfrac{1}{\pi ER^2 r\sqrt{1+k^2}}\left(1+\dfrac{4\nu r^2}{R^2}\right) \\ s_4 = \dfrac{4}{\pi ER^4 k\sqrt{1+k^2}}[1+k^2(1+\nu)]. \end{cases} \tag{31}$$

It is seen from the above solutions that the free vibrations of the spring are characterized by the wave functions $w_i(x, t)$, $i = 1, 2$. It suffices, therefore, to study the individual motion separately to obtain all the information needed and then superpose them together to get the final resultant.

The free longitudinal vibration of a rod is a classic problem. The solution can be found from almost all the vibration and applied mathematics textbooks:

$$u(x,t) = \frac{8u_0}{\pi^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}\sin\frac{(2n+1)\pi x}{2h}\cos\frac{(2n+1)\pi}{2h}\sqrt{\frac{E}{\rho}}\,t. \tag{32}$$

Similarly, for the free torsional vibration of the rod, the solution is

$$\theta(x,t) = \frac{8\theta_0}{\pi^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}\sin\frac{(2n+1)\pi x}{2h}\cos\frac{(2n+1)\pi}{2h}\sqrt{\frac{G}{\rho}}\,t. \tag{33}$$

In equations (32) and (33), ρ is the mass per unit volume, E is the Young's modulus and G is the shear modulus. The u_0 and θ_0 are initial longitudinal and rotational displacements as defined previously.

A comparison shows that the wave function $w_i(x, t)$, defined by equation (25), represents the same type of vibration mode as that of a rod except for the frequencies and amplitudes. To explore their relationships, let us consider a limiting case when the helical angle of the spring approaches 90 deg. In such a case, the shape of the spring degenerates to that of a rod, and it is expected that the spring will behave like the rod.

When the helical angle α approaches 90 deg, the coefficient k becomes larger and larger, and finally approaches infinity. If we consider, in this limiting case, that the helix radius r tends toward zero, and that the mass and mass moment of inertia per unit length of the spring become

$$\begin{cases} \gamma = \pi R^2\rho \\ \mu = I_p\rho \end{cases} \tag{34}$$

respectively, where I_p is the polar moment of inertia of the cross section, equation (2) yields

$$\begin{cases} k_1 = \pi ER^2 \\ k_2 = k_3 = 0 \\ k_4 = GI_p. \end{cases} \tag{35}$$

Then, by equation (12),

$$\begin{cases} \omega_1 = h\sqrt{\dfrac{\rho}{E}} \\ \omega_2 = h\sqrt{\dfrac{\rho}{G}}. \end{cases} \tag{36}$$

By equations (20)–(23),

$$\begin{cases} a_1 = d_2 = 1 \\ a_2 = b_1 = b_2 = c_1 = c_2 = d_1 = 0 \end{cases} \tag{37}$$

Finally, equation (24) gives

$$\begin{cases} u(x,t) = u_0 w_1(x,t) \\ \theta(x,t) = \theta_0 w_2(x,t) \end{cases} \tag{38}$$

and they are the rod solutions, equations (32) and (33). In other words, when the helical angle α approaches 90 deg, the spring does behave like a rod; initial axial displacement or force causes longitudinal vibration only, and initial rotational displacement or torque produces just torsional vibration. From equation (36), it is seen that the wave function $w_1(x, t)$ characterizes the longitudinal vibration, while $w_2(x, t)$ characterizes the torsional vibration of the spring. The spring vibration is a mix of these two waves and their interaction depends on the helical angle. For example, if there is no initial rotational displacement, the maximum possible axial displacement in the vibration, equation (24), is the initial stretch u_0. The longitudinal waves shares the part of a_1u_0, whereas the torsional wave has an amplitude of a_2u_0. When the helical angle is enlarged,the coefficient a_1 increases and approaches one, while a_2 decreases and approaches zero. That is, the influence of the torsional deformation on the extensional-compressive deformation becomes weaker, and finally the longitudinal vibration prevails.

Now let us take a closer look at the wave function itself. As has been previously shown, the wave function is the classical solution to the longitudinal vibration of a rod. However, it appears that no one before manipulated it further to obtain its closed form and thereby to reveal some detailed information about the process of the wave propagation. In fact, the wave function $w_i(x, t)$ can be expressed in a very simple closed form:

$$w_i(x,t) = \begin{cases} (-1)^m\left(\dfrac{t}{\omega_i} - 2m + 1\right) \\ \qquad \left(2m-1-\dfrac{x}{h}\right)\omega_i \le t \le \left(2m-1+\dfrac{x}{h}\right)\omega_i \\ (-1)^m\dfrac{x}{h} \\ \qquad \left(2m-1+\dfrac{x}{h}\right)\omega_i \le t \le \left(2m+1-\dfrac{x}{h}\right)\omega_i \\ \qquad m = 0,1,\ldots. \end{cases} \tag{39}$$

The wave function $w_i(x, t)$ as a function of time t for some fixed x is shown in Fig. 3(*a*). It is a very simple trapezoidal wave, and it degenerates to a triangular wave at the end $x = h$ as shown in Fig. 4. The velocities and strains are composed of periodic square waves as shown in Figs. 3(*b*) and 3(*c*).

Based on these results, we can gain a deep insight into the spring motion characterized by one of the wave functions $w_i(x, t)$. For this purpose, let us consider the displacement component

$$u_i(x,t) = u_0 a_i w_i(x,t) \tag{40}$$

of the vibration, equation (24).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

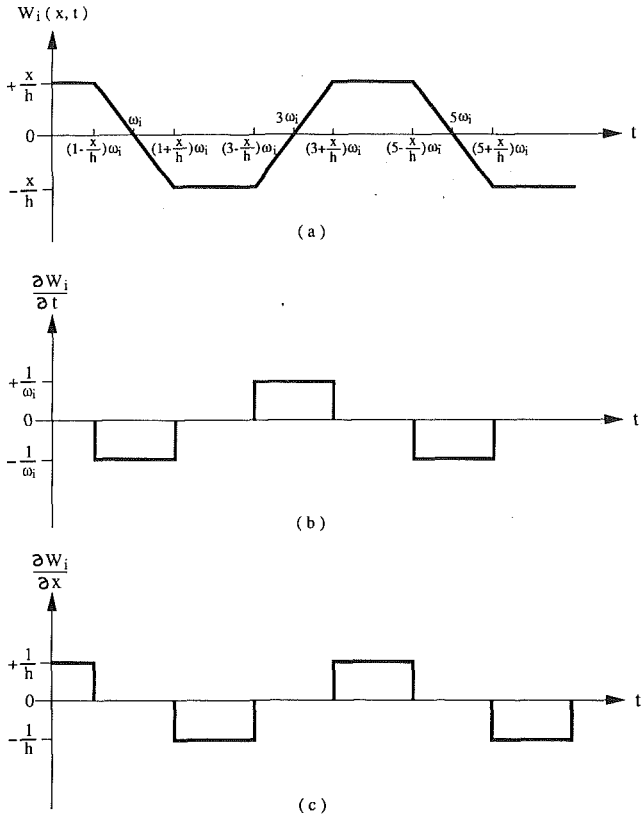


Fig. 3 Wave function and its derivatives for some fixed x

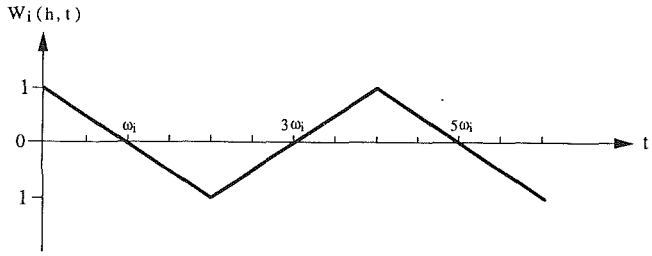


Fig. 4 Wave function at the end $x = h$

It is obvious that the motion described in the following text represents the complete longitudinal or torsional vibration of a rod, but only a component part of the vibration of a spring.

At time $t = 0$, the spring is at rest with a prestretch, which gives displacement $u = u_0 a_i$ at the end $x = h$ and $u = \frac{u_0 a_i}{h} x$ at some internal point x. As soon as the initial force is removed, the end point starts to move up towards the original equilibrium position with a velocity $\frac{u_0 a_i}{\omega_i}$, while the wave propagates with a velocity $\frac{1}{\omega_i}$. The point x, however, is still at rest, undisturbed until $t = \left(1 - \frac{x}{h}\right)\omega_i$ when the wave reaches it; then it moves up also. At time $t = \omega_i$, all the spring points contract back to their original positions, and the spring resumes it original unstretched state with $u = 0$ everywhere. But the wave pushes all the points further up, and at time $t = \left(1 + \frac{x}{h}\right)\omega_i$, the spring point x reaches its most compressive state with $u = -\frac{u_0 a_i}{h} x$. On the other hand, at $t = w_i$ when the spring resumes its unstretched state, the wave reaches the end point $x = 0$ and starts to reflect back towards the end $x = h$. At the same time $t = \left(1 + \frac{x}{h}\right)\omega_i$ when the spring point x reaches its most compressive state, the reflected wave arrives at the x position. Due to the interaction of the propagating and the reflected wave, the spring point x will remain in the most compressive state for a duration of $t = 2\left(1 - \frac{x}{h}\right)\omega_i$ as if it were kept in a "frozen state." After that, the reflected wave prevails and pushes the point x down towards its most extensive position again. The motion has a period of $4w_i$, while the frozen state time depends on the location x. At the end $x = h$, the frozen state time is zero and the motion degenerates to a triangular wave.

The velocity wave and the strain wave characterized by the wave function $w_i(x, t)$ are square as shown in Figs. 3(b) and 3(c). When the point x of the spring is under the frozen state, the velocity there is zero, whereas the strain reaches the maximum. On the other hand, when this point moves, it moves with a uniform velocity $\frac{u_0 a_i}{\omega_i}$, whereas the strain there vanishes.

For a rod, the above wave-propagating process describes the complete motion for longitudinal vibration or torsional vibration. For a spring, however, the actual displacements, velocities, and strains are the superposition of two similar waves. Due to the differences between the periods and amplitudes, the resultant motion can be very complex, as will be shown in a paper to follow.

V Conclusions

This study of the coupled extensional-torsional dynamic response of the helical springs revealed some interesting characteristics of spring behavior, which resulted in some interesting and important vibration phenomena. The research discovered that the free vibration of a rod can be described by a simple trapezoidal wave, and as such, the motion of the particles of the rod is not continuous, but "frozen" for a certain duration at their most extensional or compressive state. The vibration of the spring was found to be the resultant of two such trapezoidal waves, one representing the extensional-compressive deformation and the other one the torsional deformation, and due to the interaction between these two waves the spring vibration became complex.

References

Costello, G. A., 1975, "Radial Expansion of Impacted Helical Springs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 789–792.

Costello, G. A., 1983, "Stresses in Multilayered Cables," *ASME Journal of Energy Resources Technology*, Vol. 105, No. 3, pp. 337–340.

Curran, J. P. S., 1970, "Further Considerations in Injector Design for High Specific Output Diesel Engines," *I. Mech. E. Symp.* (Critical Factors in the Application of Diesel Engines), Southampton University.

Dick, J., 1957, "Shock-Waves in Helical Springs," *The Engineer*.

Gross, S. C., 1959, "Coil Spring Performance," *Automobile Engineering*.

Jiang, W., Jones, W. K., Wu, K. H., and Wang, T. L., 1989, "Non-Linear and Linear, Static and Dynamic Analyses of Helical Springs," *Proceedings of the 30th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamic and Materials Conference*, Mobile, Ala., Apr., pp. 386–395.

Love, A. E. H., 1899, "The Propagation of Waves of Elastic Displacement Along a Helical Wire," *Trans. Camb. Phil. Soc.*, Vol. 18, pp. 364–374.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Dover, New York.

Michell, J. H., 1890, "The Small Deformation of Curves and Surfaces with Application to the Vibrations of a Helix and a Circular Ring," *Mess. Math*, Vol. 19, pp. 68–82.

Mottershead, J. E., 1980, "Finite Elements for Dynamical Analysis of Helical Rods," *International Journal of Mechanical Science*, Vol. 22, pp. 267–283.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Pearson, D., 1988, "Dynamic Behavior of Helical Springs," *Shock and Vibration Digest*, Vol. 20, No. 7, pp. 3–9.

Pearson, D., and Wittrick, W. H., 1986, "An Exact Solution for the Vibration of Helical Springs Using a Bernoulli-Euler Model," *International Journal of Mechanical Science*, Vol. 28, No. 2, pp. 83–96.

Phillips, J. W., and Costello, G. A., 1972, "Large Deflections of Impacted Helical Springs," *Journal of the Acoustical Society of America*, Vol. 51, No. 3, Part 2, pp. 967–973.

Phillips, J. M., and Costello, G. A., 1977, "Axial Impact of Twisted Wire Cables," ASME JOURNAL OF THE APPLIED MECHANICS, Vol. 99, No. 1, pp. 127–131.

Samras, R. K., Skop, R. A., and Milburn, D. A., 1974, "An Analysis of Coupled Extensional-Torsional Oscillations in Wire Rope," *ASME Journal of Engineering for Industry*, pp. 1130–1135.

Sinha, S. K., and Costello, G. A., 1978, "The Numerical Solution of the Dynamic Response of Helical Springs," *Intl. J. Numer. Methods Engrg.*, Vol. 12, pp. 949–961.

Wittrick, W. H., 1966, "On Elastic Wave Propagation in Helical Springs," *International Journal of Mechanical Science*, Vol. 8, No. 25, pp. 25–47.

Yoshimura, Y., and Murata, Y., 1952, "On the Elastic Waves Propagated Along Coil Springs," *Rep. Inst. Sci. and Tech. Tokyo University*, Vol. 6, No. 1, pp. 27–35.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stability of Linear Conservative Gyroscopic Systems

J. A. Walker
Department of Mechanical Engineering,
Northwestern University,
Evanston, IL 60208

Sufficient conditions are obtained for the stability and instability of linear conservative gyroscopic systems. The conditions are nonspectral, involve only the definiteness of certain combinations of the coefficient matrices, and may yield useful design constraints. An example is employed to compare these results with earlier results of the same type.

1 Introduction

Systems of interest here are linear conservative gyroscopic systems described by the equation

$$M\ddot{q}(t)+\tilde{C}\dot{q}(t)+\tilde{K}q(t)=0, \tag{1}$$

where M, $\tilde{C}$, and $\tilde{K}$ are $n\times n$ constant real matrices, $\tilde{K}$ is symmetric and nonsingular, M is symmetric and positive definite ($M>0$), $\tilde{C}$ is skew ($\tilde{C}^T=-\tilde{C}$), and $q(t)$ is a real $n\times 1$ time-varying matrix. A linear system described by equation (1) is called "conservative" because the conditions $\tilde{C}^T=-\tilde{C}$ and $\tilde{K}^T=\tilde{K}$ lead to an energy conservation result; it is called "gyroscopic" because $\tilde{C}$ is assumed skew ($\tilde{C}^T=-\tilde{C}$) but *not* zero. A recent paper in the *ASME Journal of Applied Mechanics* (Inman, 1988) has stimulated new interest in the stability properties of such systems.

Equation (1) can be simplified by noting that M has a real symmetric and positive-definite square root $M^{1/2}$. Hence, we may transform the coordinates by $x(t)\equiv M^{1/2}q(t)$, define $C\equiv M^{-1/2}\tilde{C}M^{-1/2}$, $K\equiv M^{-1/2}\tilde{K}M^{-1/2}$, and consider the simplified equation

$$\ddot{x}(t)+C\dot{x}(t)+Kx(t)=0 \tag{2}$$

in place of equation (1). The assumptions on $\tilde{C}$ and $\tilde{K}$ imply that $C^T=-C$, $K^T=K$, and K is nonsingular.

As equation (1) is linear, every solution $q(t)$ of equation (1) and $x(t)$ of equation (2) has the same stability properties as the equilibrium $(x,\dot{x})=(0,0)$ of equation (2). Hence, "stability of the system" is determined by the stability of the equilibrium $(x,\dot{x})=(0,0)$ of equation (2).

If $C=0$ it is easily established by spectral analysis (eigenvalue analysis) of equation (2) that the equilibrium is stable if K is positive definite ($K>0$) and unstable otherwise. For $C^T=-C\neq 0$, the stability properties may be very different; it is quite possible that the equilibrium may be "gyroscopically stabilized" for K negative definite ($K<0$) or indefinite. Earlier results on this problem are described in the following Section 2, and new results are obtained in the subsequent Section 3.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 7, 1989; final revision, Feb. 5, 1990.

2 Earlier Stability Criteria

The stability properties of the equilibrium of equation (2) have long been of interest, and the most simple nonspectral criteria are as follows:

(A) The equilibrium is stable if $K>0$ (Lagrange-Dirichlet condition).

(B) The equilibrium is unstable if $C^TC+4K<0$ (Hagedorn, 1975).

(C) The equilibrium is stable if $CK=KC$ and $C^TC+4K>0$ (Huseyin, Hagedorn, and Teschner, 1983).

Criterion (A) is the Lagrange-Dirichlet "principle of minimum potential energy"; it requires only that total energy $(\dot{x}^T\dot{x}+x^TKx)/2$ of equation (2) have an isolated minimum at the equilibrium $(x,\dot{x})=(0,0)$, and is valid for any C such that $C^T+C\geq 0$. Here we are assuming that $C^T=-C$. Criterion (B) shows that gyroscopic stabilization cannot occur when $C^TC+4K<0$, and criterion (C) shows that gyroscopic stabilization does occur under certain other conditions.

If $CK=KC$, as assumed in criterion (C), then the system (2) may be completely decoupled by a complex coordinate transformation; this is a very restrictive condition. It was proved in (Huseyin, Hagedorn, and Teschner, 1983) that although the condition $CK=KC$ can be replaced in criterion (C) by the seemingly weaker condition $H\equiv CK-KC\geq 0$, it is pointless to do so since H is indefinite unless $H=0$.

An additional sufficient condition for stability was recently proposed in (Inman, 1988):

(D) The equilibrium is stable if $K<0$ and $C^TC+4K-\epsilon I>0$, where 2ϵ is the largest eigenvalue of C^TC.

Unfortunately, Inman's proof of (D) is erroneous. Proposition (D) may be correct, but it is unproven.

When the foregoing criteria are applied to a two degree-of-freedom system described by

$$\begin{bmatrix}1&0\\0&1\end{bmatrix}\ddot{x}(t)+\begin{bmatrix}0&-4\\4&0\end{bmatrix}\dot{x}(t)+\begin{bmatrix}k_1&0\\0&k_2\end{bmatrix}x(t)=0, \tag{3}$$

where

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

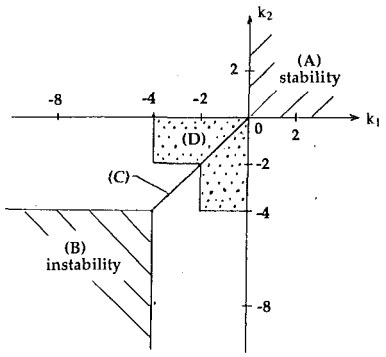


Fig. 1 Earlier criteria

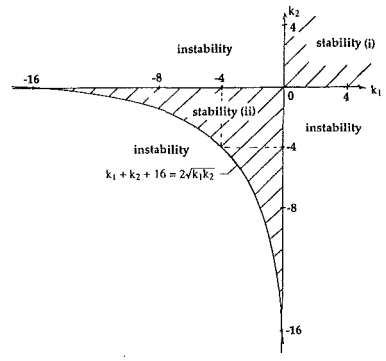


Fig. 2 Exact region of stability

$$C=\begin{bmatrix}0 & -4\\ 4 & 0\end{bmatrix}=-C^T,\quad K=\begin{bmatrix}k_1 & 0\\ 0 & k_2\end{bmatrix}=K^T,$$

they produce the results displayed in Fig. 1. Region (A) yields stability by criterion (A), and region (B) yields instability by criterion (B). The line segment (C) between (0,0) and $(-4,-4)$ yields stability by criterion (C). Proposition (D) claims that region (D) yields stability. With a shift of the parameters $(k_i \rightarrow k_i - 7)$, equation (3) is identical to the example of (Inman, 1988).

By means of spectral analysis (eigenvalue analysis), one can show that the equilibrium of equation (3) is stable if and only if $k_1k_2>0$ and $k_1+k_2+16 > 2\sqrt{k_1k_2}$. Hence, in Fig. 2 the exact region of stability consists of (i) the quadrant $k_1>0$, $k_2>0$, and (ii) that portion of the opposing quadrant $k_1<0$, $k_2<0$ such that $k_1+k_2+16 > 2\sqrt{k_1k_2}$; region (ii) is gyroscopically stabilized. The remainder of the parameter plane corresponds to instability. Proposition (D) is not contradicted by this example.

The existence of the gyroscopically stabilized region (ii) is critically dependent upon our gyroscopic assumption $C^T=-C\neq 0$, whereas region (i) remains a stability region for any C such that $C^T+C\geq 0$. Gyroscopic stabilization is an extremely delicate phenomenon, easily upset by the slightest dissipation, and often difficult to predict.

Although spectral analysis is an exact method of stability analysis for equations (1) and (2), and is easy for the example (3), it rapidly becomes impractical as the order $2n$ of equations (1) and (2) is increased. This is why one seeks simple sufficiency criteria, such as (A)-(C), which *remain* simple regardless of the order $2n$ of the system.

On the other hand, Figs. 1 and 2 for the example (3) clearly demonstrate that criteria (A)-(C), and even conjecture (D), are not very sharp. Moreover, unless (D) is true, there is *no* nonspectral sufficient condition for gyroscopic stabilization when C and K fail to commute. At the cost of introducing a scalar parameter, we obtain much sharper criteria in the following Section 3.

3 New Stability Criteria

Let G_i be any symmetric matrix which commutes with $K(G_iK=KG_i)$ and define three functions V_i of the state $(x,\dot{x})$ as follows:

$$V_1(x,\dot{x})=\dot{x}^TG_1\dot{x}+x^TG_1Kx,$$
$$V_2(x,\dot{x})=(\dot{x}+Cx)^TG_2(\dot{x}+Cx)+x^TG_2Kx,$$
$$V_3(x,\dot{x})=-4x^TG_3(\dot{x}+Cx).$$

As K is symmetric, the assumptions on G_i imply that G_iK is symmetric. Hence, along every motion of equation (2) the time derivative of V_i is given by

$$\dot{V}_1(x,\dot{x})=-2\dot{x}^TG_1C\dot{x},$$
$$\dot{V}_2(x,\dot{x})=-2x^TKG_2Cx,$$
$$\dot{V}_3(x,\dot{x})=-(2\dot{x}+Cx)^TG_3(2\dot{x}+Cx)+x^T(C^TG_3C+4G_3K)x.$$

Since $C^T=-C$, we see that V_1 is a first integral of equation (2) if G_1 commutes with C; i.e., $\dot{V}_1=0$ if $G_1C=CG_1$. We also see that V_2 is a first integral of equation (2) if G_2K commutes with C; i.e., $\dot{V}_2=0$ if $(G_2K)C=C(G_2K)$. We also see that V_3 is negative definite if $G_3>0$ and $C^TG_3C+4G_3K<0$.

These facts allow criteria (A), (B), and (C) to be established by the theorems of Liapunov (Hahn, 1963). Choosing $G_1=I$, we see that $\dot{V}_1\equiv 0$ and V_1 is positive definite if $K>0$; hence criterion (A) is established by Liapunov's stability theorem. If $C^TC+K<0$ and we choose $G_3=I$, then $\dot{V}_3$ is negative definite and V_3 is indefinite; hence criterion (B) of (Hagedorn, 1975) is established by Liapunov's first instability theorem. If $CK=KC$, $C^TC+4K>0$, and we choose $G_1=G_2=I$, then $V_1+V_2\equiv 0$ and

$$2V_1(x,\dot{x})+2V_2(x,\dot{x})=(2\dot{x}+Cx)^T(2\dot{x}+Cx)+x^T(C^TC+4K)x$$

is positive definite; hence, criterion (C) of (Huseyin, Hagedorn, and Tescher, 1983) follows from Liapunov's stability theorem. We have found no way to establish proposition (D).

Criteria (A)-(C) provide very simple sufficient conditions for stability and instability, and they involve no undetermined parameters. In order to obtain sharper criteria of the same type, it appears that one or more undetermined parameters must be introduced. The following Theorem I and Theorem II are one-parameter criteria. They are proved by making the functions V_1, V_2, and V_3 into new Liapunov functions (Walker, 1980), and they provide much better results for the example (3).

Theorem I: The equilibrium is stable if there exists a real number α such that

$$K(K-\alpha I)>0,$$
$$K-\alpha I-\alpha C^T(K-\alpha I)^{-1}C>0.$$

Theorem II. The equilibrium is unstable if there exists a real number λ such that

$$I-\lambda K>0,$$
$$C^T(I-\lambda K)^{-1}C+4K(I-\lambda K)^{-1}<0.$$

Theorem I is obtained by choosing $G_1=I$, $G_2=K^{-1}$, noting that $\dot{V}_1=0=\dot{V}_2$, and employing Liapunov's stability theorem with the function

$$V_1(x,\dot{x})-\alpha V_2(x,\dot{x})=z^TK^{-1}(K-\alpha I)z+x^T[K-\alpha I-\alpha C^T(K-\alpha I)^{-1}C]x,$$

where $z\equiv\dot{x}-\alpha(K-\alpha I)^{-1}Cx$. Theorem II is obtained by choosing $G_3=(I-\lambda K)^{-1}$ and employing Liapunov's first instability theorem with V_3.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

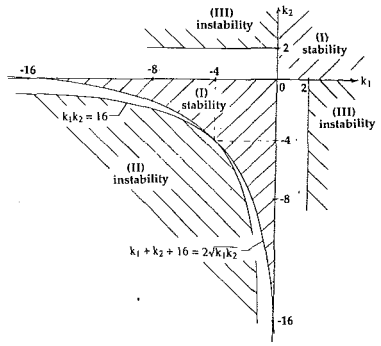


Fig. 3 New criteria

Criterion (A) follows from the choice $\alpha = 0$ in Theorem I, and criterion (B) follows from the choice $\lambda = 0$ in Theorem II. Hence, Theorems I–II are at least as good as criteria (A)–(B).

Returning to the example (3), we find that

$$I - \lambda K = \begin{bmatrix} 1-\lambda k_1 & 0 \\ 0 & 1-\lambda k_2 \end{bmatrix},$$

$$C^T(I-\lambda K)^{-1}C + 4K(I-\lambda K)^{-1} =$$

$$\beta \begin{bmatrix} 16+4k_1-\lambda k_1(16+4k_2) & 0 \\ 0 & 16+4k_1-\lambda k_2(16+4k_2) \end{bmatrix},$$

where $\beta \equiv 1/(1-\lambda k_1)(1-\lambda k_2)$. Hence, Theorem II implies instability if there exists λ such that

$$-k_1(1-\lambda k_2) > 4(1-\lambda k_1) > 0,$$
$$-k_2(1-\lambda k_1) > 4(1-\lambda k_2) > 0.$$

Such λ exists if and only if $k_1 < 0$, $k_2 < 0$, and $k_1k_2 > 16$. Therefore, Theorem II yields the instability region (II) shown in Fig. 3, and whose boundary is the hyperbola $k_1k_2 = 16$. Region (II) contains region (B) and is contained within the exact instability region. See Figs. 1, 2, and 3.

For the example (3) we also obtain

$$K - \alpha I = \begin{bmatrix} k_1-\alpha & 0 \\ 0 & k_2-\alpha \end{bmatrix},$$

$$K(K-\alpha I) = \begin{bmatrix} k_1^2-\alpha k_1 & 0 \\ 0 & k_2^2-\alpha k_2 \end{bmatrix},$$

$$K - \alpha I - \alpha C^T(K-\alpha I)^{-1}C$$

$$= \frac{1}{\gamma}\begin{bmatrix} \gamma - 16\alpha(k_1-\alpha) & 0 \\ 0 & \gamma - 16\alpha(k_2-\alpha) \end{bmatrix},$$

where $\gamma \equiv (k_1-\alpha)(k_2-\alpha)$. If $k_1 > 0$ and $k_2 > 0$, we see that the conditions of Theorem I are met for $\alpha = 0$. For $k_1 < 0$ and $k_2 < 0$, Theorem I still implies stability if there exists α such that $\alpha > k_1$, $\alpha > k_2$, and

$$\alpha^2 - \alpha(k_1+k_2+16) + k_1k_2 < 0.$$

Such α exists if and only if $k_1 + k_2 + 16 > 2\sqrt{k_1k_2}$. Thus Theorem I yields the two-part stability region (I) shown in Fig. 3. Region (I) contains regions (A) and (C), and is identical to the exact region of stability. (See Figs. 1, 2, and 3.)

Although Theorem II is clearly superior to the instability criterion (B), it requires K to be negative definite. For indefinite K, nonspectral instability criteria are quite complicated if $C + C^T = 0$, as here. The best instability criterion seems to be Corollary 3.2 of (Walker, 1988)[1], but that is a multiparameter result. We seek a one-parameter criterion for instability with K indefinite.

Spectral theory shows that any nonsingular symmetric matrix K may be uniquely decomposed into two semi-definite symmetric matrices as follows:

$$K = K_p + K_n,\ K_pK_n = 0,\ K_p \geq 0,\ K_n \leq 0. \tag{4}$$

If $K > 0$, then $K_p = K$ and $K_n = 0$. If $K < 0$, then $K_n = K$ and $K_p = 0$. If K is indefinite, then $K_p \neq 0 \neq K_n$. Defining $J \equiv K_nK^{-1}$ we see that $K_n = JK$, $K_p = (I-J)K$, J is a symmetric projection ($J^2 = J = J^T$) and $J \geq 0$.

We notice that the decomposition (4) is obvious if K is diagonal; then both K_p and K_n are diagonal. Therefore, the decomposition (4) of K is no more difficult than diagonalization of K (i.e., modal analysis). Although diagonalization of K may be laborious, it is much easier than spectral analysis of equation (2).

Once the decomposition (4) of K is known, the following instability result becomes useful:

Theorem III: The equilibrium is unstable if K is indefinite and there exists a real number $\nu > 0$ such that, with $J \equiv K_nK^{-1}$ in (4),

$$4(\nu I - J)K - C^TJ(\nu I + J)^{-1}JC \geq 0.$$

To establish Theorem III we choose $G_1 = 4\nu I$ and $G_3 = J$. Clearly, $\nu I + J > 0$, and we find

$$\dot{V}_3(x,\dot{x}) - V_1(x,\dot{x}) = -z^T(\nu I + J)z - x^T[4(\nu I - J)K - C^TJ(\nu I + J)^{-1}JC]x \leq 0,$$

where $z \equiv 2x + (\nu I + J)^{-1}JCx$. Hence, $\dot{V}_3(x,\dot{x}) < 0$ wherever $V_1(x,\dot{x}) < 0$. As K is indefinite and $\nu > 0$, every neighborhood of the equilibrium contains such points, and instability is a consequence of LaSalle's Invariance Principle (Walker, 1980)[2]. This proof shows that Theorem III is valid for negative definite K as well as indefinite K. However, $K < 0$ implies $J = I$, and the ensuing result of Theorem III is found to be identical to criterion (B).

Theorem III is quite easy to use on the example (3) where K is diagonal. If $k_1 > 0$ and $k_2 < 0$, then K is indefinite, the decomposition (4) is obvious,

$$K_p = \begin{bmatrix} k_1 & 0 \\ 0 & 0 \end{bmatrix},\ K_n = \begin{bmatrix} 0 & 0 \\ 0 & k_2 \end{bmatrix},$$

and

$$J = K_nK^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},\ \nu I + J = \begin{bmatrix} \nu & 0 \\ 0 & \nu+1 \end{bmatrix},$$

$$4(\nu I - J)K - C^TJ(\nu I + J)^{-1}JC$$

$$= \frac{4}{\nu+1}\begin{bmatrix} (\nu+1)k_1 - 4 & 0 \\ 0 & (\nu^2-1)k_2 \end{bmatrix}.$$

The conditions of Theorem III are met if there exists $\nu > 0$ such that $\nu \leq 1$ and $(\nu+1)k_1 \geq 4$; that is, Theorem III implies instability for $k_1 \geq 2$, $k_2 < 0$. By reversing our definitions of K_p and K_n, we find that Theorem III implies instability for $k_2 \geq 2$, $k_1 < 0$. Hence, Theorem III yields the two-part instability region (III) depicted in Fig. 3. Compare Figs. 1, 2, and 3.

[1]Corollary 3.2 of (Walker, 1988) establishes Theorem II. To see this set $B = 0$, $M = I$ and choose $\nu = 0$, $S = 0$, $2\Gamma = (I - \lambda K)^{-1}$ in that earlier result.

[2]If the positive semi-definiteness condition (≥ 0) in Theorem III is strengthened to definiteness (> 0), then Theorem III is another consequence of Corollary 3.2 of (Walker, 1988). To see this, set $B = 0$, $M = I$, and choose $S = 0$, $2\Gamma = J$ in that earlier result. The results of (Walker, 1988) are not restricted to linear systems, which is why Theorem III does not follow completely from (Walker, 1988).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

4 Concluding Remarks

We have been interested in providing simple stability criteria for linear conservative gyroscopic systems, criteria that involve only the definiteness of certain combinations of the coefficient matrices in equation (2). Our primary objective has been to avoid spectral analysis of equations (1) or (2), which becomes extremely difficult as the order $2n$ of (1) and (2) increases.

It may be that criteria (A)–(C) are the best possible zero-parameter criteria. We utilized the functions V_1, V_2, and V_3 of Section 3 to develop superior criteria (Theorems I–III) which involve a scalar parameter. Even better criteria could be developed in the same way by making multi-parameter choices of G_1, G_2, G_3, but this would produce complicated criteria. The accuracy of multi-parameter stability criteria can be made to approach that of spectral analysis, but so does the difficulty.

When the coefficient matrices $K = K^T$ and $C = -C^T$ are fully specified, we suggest the following simplified approach to stability analysis: First determine whether any of the unparameterized criteria (A)–(C) apply. If not, make random selections of α, λ, (if $K < 0$), or ν (if K is indefinite) and determine whether any of Theorems I–III apply. If not, try new selections of α, λ, or ν. This hit-or-miss method is easily implemented on a computer. Resort to spectral analysis of equation (2) only if no suitable α, λ, or ν is found.

Comparing Figs. 1 and 3, we see that Theorems I–III are superior to criteria (A)–(C) at estimating the true regions (Fig. 2) of stability and instability for the example (3). However, the fact that Theorem I happens to give the exact stability region for the example (3) should not be taken too seriously. It is very unlikely that any one-parameter criterion could be exact for the entire class of systems considered, and Theorem I is not exact for $n \geq 4$. The purpose of nonspectral criteria is to provide simple sufficient conditions for stability or instability, and this is all that has been achieved here.

References

Hagedorn, P., 1975, "Uber Die Instabilitat Konservativer Systeme Mit Gyroskopischen Kraften," *Achieve for Rational Mechanics and Analysis*, Vol. 58, pp. 1–9.

Hahn, W., 1963, *Theory and Application of Lapunov's Direct Method*, Prentice-Hall International, London.

Huseyin, K., Hagedorn, P., and Teschner, W., 1983, "On the Stability of Linear Conservative Gyroscopic Systems," *Journal of Applied Mathematics and Physics* (ZAMP), Vol. 34, pp. 807–815.

Inman, D., 1988, "A Sufficient Condition for the Stability of Conservative Gyroscopic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 895–898.

Walker, J., 1980, *Dynamical Systems and Evolution Equations*, Plenum Press, New York.

Walker, J., 1988, "Pseudodissipative Systems I: Stability of Generalized Equilibria," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 681–686

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Y. M. Ram

S. G. Braun
Mem. ASME

Department of Mechanical Engineering,
Technion-Israel Institute of Technology,
Haifa 32000, Israel

An Inverse Problem Associated With Modification of Incomplete Dynamic Systems

Given an incomplete set of natural frequencies and modes of a linear structure from a vibration test, the problem of computing the necessary changes in the structure to reach a desired spectra is considered. A criterion of optimality which gives the best approximation from a certain subspace is formulated. Then formulae for evaluating the family of all mathematical solutions are given, resulting in a class of mathematical incremental mass and stiffness matrices. Finally, we describe the problem of the realization of the mathematical solutions by constructing a model of a given type.

1 Introduction

This paper addresses one aspect of structural modification, designated here as the "Inverse Problem." In this problem we search for a modification in the structure which will result in a specified new set of natural frequencies and mode shapes.

By experimental modal analysis it is possible to obtain an incomplete modal model of the structure. The task in this paper is to derive an optimal approximation for the incremental mass and stiffness matrices needed to solve the inverse problem based on truncated modal model. The contributions of this paper are: (a) defining an optimization problem needed to solve the inverse modification problem based on incomplete modal model, and (b) the evaluation of the class of modification matrices which form the mathematical solutions of the inverse problem. The solution presented here is based on the optimal solution for the direct problem, which we have already derived elsewhere (Ram et al., 1988).

The second important task is to identify the constructable solutions, i.e., the problem of determining the solutions which can be realized by a physical change in the structure. This problem is discussed, but we do not give an efficient solution to this problem here.

The paper is organized as follows: The general problem and the necessary background are given in Section 2. The "Direct Modification Problem" is reviewed in Section 3, followed by the mathematical formulation of the "Inverse Problem" in Section 4. Section 5 describes an intermediate step needed to obtain some of the mode shapes of the modified structure. A class of mathematical solutions is given in Section 6. The reconstruction problem is discussed in Section 7, and a numerical example is given in Section 8.

2 Background

Consider a linear m degrees-of-freedom system which is characterized by the following generalized eigenvalue problem

$$K\Phi = M\Phi\Lambda;\ M,K,\Phi\epsilon R^{m\times m},\ \Lambda = \mathrm{DIAG}\{\lambda_i;\ i=1,\ \ldots,\ m\} \quad (1)$$

where K and M are the stiffness and the mass matrix, respectively. The mass and stiffness matrices are symmetric, M is positive definite, and K is a non-negative definite matrix. The mass normalized modal matrix Φ satisfies

$$\Phi^T M\Phi = I_m. \quad (2)$$

Partition Φ and Λ is in the form

$$\Phi = [\Phi_1 \mid \Phi_2] \quad (3)$$

$$\Lambda = \begin{bmatrix} \Lambda_1 & \vdots & 0 \\ \text{----} & \vdots & \text{----} \\ 0 & \vdots & \Lambda_2 \end{bmatrix} \quad (4)$$

where Φ_1 is an m-by-n matrix and $\Lambda_1 = \mathrm{DIAG}\ \{\lambda_i;\ i=1,\ \ldots,\ n\}$. We refer to the system (1) as the *original system*. The original system is subject to changes. Suppose that the *modified system* is characterized by the following eigenvalue problem:

$$(K+\delta K)\tilde{\Phi} = (M+\delta M)\tilde{\Phi}\tilde{\Lambda}, \quad (5)$$

where δK and δM are *incremental matrices* due to the mass and the stiffness modification. The matrices $K+\delta K$ and $M+\delta M$ are semi-definite and positive definite symmetric matrices, respectively. In some applications, due to the complexity of the structure, no reasonable analytical model of the mass and stiffness matrices of the original system can be evaluated. If the original system already exists, then a vibration test is a possible source of information concerning the dynamic behavior of the structure. Modal testing results provide an in-

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 16, 1989; final revision, May 16, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

complete set of natural frequencies and modes corresponding to the low frequencies range. Let Φ_1 and Λ_1 be the truncated modal testing results.

There are two major problems associated with structural modification. We refer to the first problem as the *direct problem*. In this problem the following are assumed to be known: (a) modal testing results of the original system (Φ_1 and Λ_1), and (b) the design incremental matrices (δM and δK). The objective of the direct problem is to predict some of the natural frequencies and modes of the modified system based on the given data. The matrices $\tilde{\Phi}$ and $\tilde{\Lambda}$ (see equation (5)) construct the exact eigensolution of the modified structure. Clearly, due to the modal truncation, the given data do not provide the sufficient information needed to evaluate the exact solution of the problem. In many publications the direct problem is considered (see, e,g., Elliott and Mitchell 1987; Ewins, 1986). The usual way to tackle the direct problem is to express the eigensolution of the modified system in terms of modal summation formulae (usually in matrix form). When the complete spectral decomposition of the original system is given (i.e., the complete set of eigenpairs of the original system is known), the modal summation procedure yields to the exact eigensolution of the modified system. In practice, where only an incomplete set of eigenpairs can be identified, the summation is taken over the truncated set of eigenpairs, resulting in a *truncation error*. It is believed that modal truncation is blamed for causing poor modification results (Elliott and Mitchell, 1987). We recently tackled the truncation problem in a series of papers. We developed (Ram et al., 1990a) upper and lower bounds for the natural frequencies of the modified structure for the case where the mass matrix of the original system is known and where the modifications effect only the stiffness matrix. Since it is hard to think of a realistic modification which changes the stiffness matrix but leaves the mass matrix unaltered, the results in Ram et al. (1990a) are mainly of theoretical interest. We have augmented (Ram and Braun, 1990b) the results to the case where the mass and stiffness matrices of the original system are subject to changes, but the mass matrix of the original system is still known. A slight modification in the algorithm for determining the lower bounds for the truncation error is given in Ram and Braun (1990c) resulting in improved bounds which are the best possible bounds for the given information.

The second problem is the *inverse problem*. In this problem a truncated set of natural frequencies and their associated mode shapes of the original structure is given based on a vibration test. The goal is to define the necessary incremental mass and stiffness matrices to get a desired spectra. An excellent source which covers many aspects of inverse problems in vibration is the book of Gladwell (1986). Due to the truncation, our problem is more complicated than the classical inverse problems appearing in this book. Zhang et al. (1989) have developed a procedure for determining the amount of mass modifications to the structure for certain dynamic requirements. Their approach is based on perturbation theory, and thus holds only for small changes in the mass matrix.

As in the direct problem, the given incomplete data provides us with insufficient information to derive the exact solution. In order to circumvent this difficulty, we define a *residual error matrix*. Then, we find the family of all incremental matrices which minimize the Frobenious norm of the residual matrix from a certain subspace. The resulting solutions are optimal in the sense of the Rayleigh-Ritz idea. These solutions are designated as *mathematical solutions*. The problem of reconstructing realistic incremental mass and stiffness matrices is also introduced. For this purpose we define a family of *physical incremental matrices* as a function of realistic changes in the geometrical quantities (such as changing the cross-sectional areas or the cross-sectional moments of inertia) and in the material properties (material density and Young's modulus of elasticity) of the structure. In general, the family of mathematical solutions does not intersect the family of physical solutions. An optimization problem which chooses a particular realistic solution, which is as close as possible to the mathematical solution, is derived resulting in a nonlinear optimization problem. It is beyond the scope of this paper to present a stable algorithm for evaluating the solution of the last optimization problem.

3 The Direct Problem

The objective of the direct problem is to predict some of the eigenpairs of the modified system based on the knowledge of an incomplete set of natural frequencies and modes of the unmodified system and on the incremental matrices due to the modification. Let $\tilde{\Phi}_1$ be the m-by-n submatrix of $\tilde{\Phi}$, where $\tilde{\Phi} = [\tilde{\Phi}_1, \tilde{\Phi}_2]$; and let $\tilde{\Lambda}_1$ be the n-by-n leading principal submatrix of $\tilde{\Lambda}$. Then the direct problem is: *to find an approximation for $\tilde{\Lambda}_1$ and $\tilde{\Phi}_1$ where Λ_1, Φ_1, δM, and δK are given.*

Consider now the following eigenvalue problem:

$$(\Lambda_1 + \Phi_1^T \delta K \Phi_1) W = (I_n + \Phi_1^T \delta M \Phi_1) W\Theta;\ W^T (I_n + \Phi_1^T \delta M \Phi_1) W = I_n \qquad (6)$$

where $\Theta = \text{DIAG}\ \{\theta_i;\ i=1, \ldots, n\}$ and $W \epsilon R^{n \times n}$.

It has been shown (Ram et al., 1988; Ram and Braun, 1989) that the Rayleigh-Ritz solution of the direct modification problem is the pair of matrices θ and $\Gamma \equiv \Phi_1 W$. This solution is the best approximation from SPAN$\{\Phi_1\}$ since $G = \Theta$ and $Z = \Gamma$ are the solution of the following minimization problem:

The Direct Problem: Find $G = \text{DIAG}\ \{g_i, i=1, \ldots, n\}$ and $Z \epsilon R^{m \times n}$ which minimize $\|(M+\delta M)^{-1/2}\ [(K+\delta K)Z - (M+\delta M)ZG]\|_F$ subject to $Z^T(M+\delta M)Z = I_n$ and $Z \epsilon$ SPAN$\{\Phi_1\}$.

The diagonal elements of Θ and the columns of Γ are the Ritz values and the *Ritz vectors* of $(K+\delta K) - \lambda(M+\delta M)$ from SPAN$\{\Phi_1\}$, respectively. It is well known that the ith smallest Ritz value is an upper bound on the ith smallest eigenvalue of $(K+\delta K) - \lambda(M+\delta M)$. Lower bounds for the eigenvalues of $(K+\delta K)\ \lambda(M+\delta M)$ based on truncated modal testing results and on the incremental matrices can be found in Ram and Braun (1990b; 1990c).

4 The Inverse Problem: Formulation

Based on our solution for the direct problem which has been briefly described in Section 3, it is now possible to define and solve the inverse problem. In the inverse problem the subject is to find incremental matrices, δM and δK, and to get a desired Ritz approximation for the eigenpairs of $(K+\delta K) - \lambda(M+\delta M)$ from SPAN$\{\Phi_1\}$. We thus define the inverse problem as follows:

The Inverse Problem: Given Φ_1, Λ_1, Θ and $\Gamma \epsilon$SPAN$\{\Phi_1\}$, find δM and δK which satisfy $(\Lambda_1 + \Phi_1^T \delta K \Phi_1) W = (I_n + \Phi_1^T \delta M \Phi_1) W\Theta$, where $\Gamma = \Phi_1 W$.

An additional discussion is needed concerning the requirement $\Gamma \epsilon$SPAN$\{\Phi_1\}$. In Section 3 where the direct problem is reviewed, a Rayleigh-Ritz approximation from SPAN$\{\Phi_1\}$ for eigenpairs of $(K+\delta K) - \lambda(M+\delta M)$ is presented. The use of the particular subspace which is spanned by the columns of the matrix Φ_1 enable one to substitute the bi-orthogonal relations $\Phi_1^T M \Phi_1 = I_n$ and $\Phi_1^T K \Phi_1 = \Lambda_1$ in the solution of the direct problem. This circumvents the difficulty that neither M nor K are known. Thus, the requirement that $\Gamma \epsilon$SPAN$\{\Phi_1\}$ is essential.[1] The next section presents a systematic way to obtain such a matrix Γ, while the mathematical solution for the inverse problem is given in Section 6.

[1] More details on this issue can be found in Ram and Braun (1989).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

5 Evaluating the Mode Shapes

In the last section we assumed that Θ and Γ are given, where $\Gamma \epsilon \text{SPAN}\{\Phi_1\}$. The diagonal matrix θ is determined by the desired location of the natural frequencies of the modified system. It is also logical to assume that an *ad hoc* matrix $P \epsilon R^{m \times n}$ can be obtained by the designer, where the columns of P are the desired modes of the modified system. But, in general, P is not in the columns range of Φ_1. So, we shall look for the matrix Γ from span $\{\Phi_1\}$ which is as close as possible to P, in the sense of minimizing a certain matrix norm. However, frequently the designer wishes to satisfy several requirements concerning the mode shapes of the modified structure, such as the determination of the location of the nodal points of a particular mode. Thus, in the following, several constraints are added. We begin the analysis with the statement that if, and only if, $\Gamma \epsilon \text{SPAN}\{\Phi_1\}$ then there exist an n-by-n matrix W such that

$$\Gamma = \Phi_1 W. \tag{7}$$

So, in order to obtain the matrix Γ it is enough to find the matrix W and then to multiply it by Φ_1.

Let $W = [w_1 \mid w_2 \mid \ldots \mid w_n]$ and $P = [p_1 \mid p_2 \mid \ldots \mid p_n]$ be the column partitioning of W and P, respectively. In order to specify the constraints associated with the ith mode of the modified system, we define the matrix $C^i \epsilon R^{q \times m}$ and the vector $b^i \epsilon R^q$ where $q \leq n$ is the number of constraints. The linear equality constraints are of the form

$$C^i \Phi_1 w_i = b^i;\ i = 1, \ldots, n. \tag{8}$$

We shall next describe two kinds of constraints.

(*a*) If the kth ($1 \leq k \leq q$) constraint associated with the ith mode of the modified system is $\Gamma_{ji} = \alpha$, then the kth row of C^i and the kth component of b_i are defined as follows:

$$C^i_{ks} = \begin{cases} 0, & s \neq j \\ 1, & s = j \end{cases};\ s = 1, 2, \ldots, m \tag{9}$$

$$b^i_k = \alpha.$$

(*b*) If the k ($1 \leq k \leq q$) constraint associated with the ith mode of the modified system is $\Gamma_{ji}/\Gamma_{ri} = \beta$, then

$$C^i_{ks} = \begin{cases} 0, & s \neq j \text{ and } s \neq r \\ 1, & s = j \\ -\beta, & s = r \end{cases};\ s = 1, 2, \ldots, m \tag{10}$$

$$b^i_k = 0.$$

The following problem determines each of the columns of the matrix W:

For $i = 1, \ldots, n$ find w_i which minimize

$$\|\Phi_1 w_i - p_i\|_2$$

subject to (11)

$$C^i \Phi_1 w_i = b^i.$$

The least square problem with linear equality constraints has a closed-form solution (Golub and Van Loan, 1983, pp. 405–414). A possible procedure for determining w_i, $i = 1, \ldots, n$; consists on the following steps:

Step 1 Compute the QR decomposition of $(C^i \Phi_1)^T = Q \begin{bmatrix} R \\ --- \\ 0 \end{bmatrix}$ where Q is an n-by-n orthonormal matrix and R is a q-by-q upper triangular matrix.

Step 2 Set the submatrices $A_1 \epsilon R^{m \times q}$ and $A_2 \epsilon R^{m \times n-q}$ by $[A_1 \mid A_2] = \Phi_1 Q$.

Step 3 By solving $R^T y = b^i$ obtain $y \epsilon R^q$.

Step 4 Evaluate $w_i = Q \begin{bmatrix} y \\ \hline A_2^+ (p_i - A_1 y) \end{bmatrix}$ where A_2^+ is the Moore-Penrose pseudo-inverse of A_2.

After evaluating w_i; $i = 1, \ldots, n$; the matrix Γ can be found by using equation (7). It follows from equation (7) that $\Gamma \epsilon \text{SPAN}\{\Phi_1\}$. If $C^i \Phi_1 \epsilon$ RANGE(Φ_1) then the ith Ritz vector $\Phi_1 w_i$ satisfies the equality constraints, if $C^i \Phi_1$ is not in the columns range of Φ_1 then $\Phi_1 w_i$ is as close as possible to satisfy the ith constraint. The matrix Γ is an optimal approximation for P, since $\|\Gamma - P\|_F$ is minimized subject to the linear equality constraints. For the sake of demonstration, consider the following example:

Example.

Let

$$\Phi_1 = \begin{bmatrix} 2/3 & -\sqrt{2}/2 \\ 2/3 & 0.0 \\ 2/3 & \sqrt{2}/2 \end{bmatrix} \text{ and } P = \begin{bmatrix} 0 & -\sqrt{2}/2 \\ 1/3 & 0.0 \\ 1 & \sqrt{2}/2 \end{bmatrix}.$$

Suppose we wish that the first point of the first mode of the modified system will be a nodal point; then our objective is to find a 2-by-2 matrix W which minimizes $\|\Phi_1 W - P\|_F$ subject to $\Gamma_{11} = 0$, where $\Gamma = \Phi_1 W$.

Following our notation we have to define a 1-by-3 matrix C^1 and a 1-by-1 vector b^1 associated with the linear equality constraint of the first mode. By using equation (9) we find $C^1 = [1\ 0\ 0]$ and $b^1 = [0]$. The vector w_1 which minimizes $\|\Phi_1 w_1 - p_1\|_F$ subject to $C^1 \Phi_1 w_1 = b^1$ is $w_1^T = [7/10\ \ 7\sqrt{2}/15]$.

We did not define any constraint associated with the second mode; and since $p_2 \epsilon \text{SPAN}\{\Phi_1\}$, we immediately have $w_2^T = [0\ 1]$. Therefore,

$$W = [w_1,\ w_2] = \begin{bmatrix} 7/10 & 0 \\ 7\sqrt{2}/15 & 1 \end{bmatrix}.$$

As a check

$$\Gamma = \Phi_1 W = \begin{bmatrix} 0 & -\sqrt{2}/2 \\ 7/15 & 0.0 \\ 14/15 & \sqrt{2}/2 \end{bmatrix}.$$

Note that Γ is sufficiently close to the matrix P, and indeed $\Gamma_{11} = 0$.

6 Mathematical Solution of the Inverse Problem

In this section the characterization of all mathematical solutions of the inverse problem is given. As a result, a family of modification matrices are derived. However, the resulting solutions do not necessarily correspond to realistic systems.

Premultiplying equation (6) by W^T, one gets

$$W^T (\Lambda_1 + \Phi_1^T \delta K \Phi_1) W = W^T (I_n + \Phi_1^T \delta M \Phi_1) W \Theta. \tag{12}$$

But

$$W^T (I_n + \Phi_1^T \delta M \Phi_1) W = I_n. \tag{13}$$

Thus, in order to satisfy equation (12), the following relation must hold:

$$W^T (\Lambda_1 + \Phi_1^T \delta K \Phi_1) W = \Theta. \tag{14}$$

Therefore, our objective is to find symmetric δM and δK such that

$$\Phi_1^T \delta M \Phi_1 = W^{-T} W^{-1} - I_n \tag{15}$$

and

$$\Phi_1^T \delta K \Phi_1 = W^{-T} \Theta W^{-1} - \Lambda_1. \tag{16}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

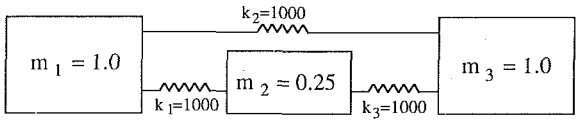


Fig. 1 The three degree-of-freedom system

Equations (15) and (16) are consistent. Moreover, the solutions of these equations are not unique. The most general symmetric solutions are (Ben-Israel and Greville, 1974, pp. 39–40)

$$\delta M = \Phi_1^{T+}(W^{-T}W^{-1} - I_n)\Phi_1^{+} + Y - \Phi_1^{T+}\Phi_1^{T}Y\Phi_1\Phi_1^{+} \quad (17)$$

and

$$\delta K = \Phi_1^{T+}(W^{-T}\Theta W^{-1} - \Lambda_1)\Phi_1^{+} + Z - \Phi_1^{T+}\Phi_1^{T}Z\Phi_1\Phi_1^{+} \quad (18)$$

where Φ_1^{+} denotes the Moore-Penrose pseudo-inverse of Φ_1 and where Y and Z are arbitrary symmetric matrices.

Equations (17) and (18) present the class of mathematical solutions of the inverse problem.

7 Reconstructing a Realizable System

So far we have developed a family of modification matrices which form the solutions to the inverse problem, but we did not guarantee that it is possible to realize even one of the solutions obtained by equations (17) and (18). In other words, the mathematical solutions did not give us a hint on how to change the geometry or the material properties of the structure to get the desired modification matrices. Furthermore, even if we know how to realize δM and δK, the realization may not be satisfactory since the locations of the modifications and the maximal amount of modification in each particular point are subject to the designer decision.

In order to tackle the realization problem let us denote the mathematical solutions of equations (17) and (18) by $\delta M_{\text{mat}}(Y)$ and $\delta K_{\text{mat}}(Z)$, respectively. Similarly, we denote the class of the physical modification matrices which can be reconstructed by realistic systems by δM_{phy} and δK_{phy}. Our motivation is to describe δM_{phy} and δK_{phy} by a finite element representation. Thus, δM_{phy} and δK_{phy} can be written in the following form

$$\delta M_{\text{phy}} = \sum_{i=1}^{m} \rho A_i B\{i\}^{T} \mathrm{M}_e \mathrm{B}\{i\} \quad (19)$$

and

$$\delta K_{\text{phy}} = \sum_{i=1}^{m} EI_i B\{i\}^{T} K_e B\{i\} \quad (20)$$

where $B\{i\}$, $i=1, \ldots, n$; are the Boolean mapping from the ith element matrix to the global matrix (Fried, 1979, pp. 55–56). EI_i and ρA_i are the change in the flexural rigidity and the change in the mass per unit length of the ith element, respectively. M_e and K_e are the mass and stiffness matrices of the element with unit characterization of flexural rigidity and mass per unit length (e.g., M_e and K_e are determined only by the trial functions and by the dimensions of the element).

The physical modification matrices can be found by solving the following minimization problem:

The Realization Problem: Find EI_i, ρA_i, $i, \ldots, m$ and symmetric matrices Y and Z which minimize $\|\delta K_{\text{phy}} - \delta K_{\text{mat}}\|$ and $\|\delta M_{\text{phy}} - \delta M_{\text{mat}}\|$ subject to inequality constraints on EI_i, ρA_i, $i=1, \ldots, m$.

The inequality constraints in the last problem are subject to the designer's decision.

It is possible to solve this problem by standard optimization techniques. It is beyond the scope of this paper to develop an efficient algorithm for the solution of the last minimization problem.

Table 1 Comparison between the design eigenpairs and the computed eigenpairs of $(K+\delta K) - \lambda(M+\delta M)$. The incremental matrices δM and δK are from equation (22).

Eigenvalue Number 1		Eigenvalue Number 2	
Desired	Computed	Desired	Computed
1200	1189.69	3500	3500.00

Mode Shape Number 1		Mode Shape Number 2	
Desired	Computed	Desired	Computed
0.0	0.01226	−0.70710	−0.70710
0.333	0.43784	0.0	0.0
1.0	0.93699	0.70710	0.70710

8 Example

Consider the three degrees-of-freedom system which is shown in Fig. 1. The two smallest eigenvalues and their corresponding mass normalized eigenvectors are:

$$\Lambda_1 = \begin{bmatrix} 0 & 0 \\ 0 & 3000 \end{bmatrix} \quad \Phi_1 = \begin{bmatrix} 2/3 & -\sqrt{2}/2 \\ 2/3 & 0.0 \\ 2/3 & \sqrt{2}/2 \end{bmatrix}.$$

Based on the knowledge of the incomplete set of eigenpairs (Λ_1 and Φ_1), we want to change the dynamic behavior of the structure by means of structural modification. The desired eigenvalues of the modified system are $\theta_1 = 1200$ and $\theta_2 = 3500$, and their corresponding eigenvectors are the columns of P, where

$$P = \begin{bmatrix} 0 & -\sqrt{2}/2 \\ 1/3 & 0.0 \\ 1 & \sqrt{2}/2 \end{bmatrix}.$$

Suppose that the linear equality constraint is $\Gamma_{11} = 0$. By the example of Section 5 we have

$$W = \begin{bmatrix} 7/10 & 0 \\ 7\sqrt{2}/15 & 1 \end{bmatrix}.$$

The family of the mathematical approximation for the mass and stiffness modification matrices can be evaluated by equations (17) and (18). For example, if we choose $Y = I_3$ and $Z = 10000 * I_3$ we get

$$\delta M = \begin{bmatrix} 1.315760 & 0.482426 & 0.649093 \\ 0.482426 & 1.149093 & -0.184240 \\ 0.649093 & -0.184240 & -0.017574 \end{bmatrix}$$

$$\delta K = \begin{bmatrix} 5640.022 & -776.643 & 2806.689 \\ -776.643 & 8056.689 & -3109.977 \\ 2806.689 & -3109.977 & 973.356 \end{bmatrix}. \quad (22)$$

For comparison the eigenvalues and the mass normalized eigenvectors of the modified system in this case are given in Table 1.

9 Conclusion

A method for approximating the necessary changes in the mass and stiffness matrices of a linear system to reach a desired dynamic behavior based on incomplete modal model of the

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

unmodified system has been developed. The mathematical formulation of the class of all mass and stiffness modification matrices has been derived. The resulting solutions are optimal in the sense of minimizing a certain residual error norm, and they are the best approximations of the desired eigenpairs from the subspace which is spanned by the given eigenvectors of the unmodified system.

The problem of choosing physically realizable changes, which are close to any of the mathematical solutions, has been formulated. There still remains the problem of developing an efficient algorithm for the evaluation of such a realizable solution which we have not treated at all in this paper.

References

Ben-Israel, A., and Greville, T. N. E., 1974, *Generalized Inverses: Theory and Applications*, John Wiley and Sons, New York.

Elliott, D. J., and Mitchell, L. D., 1987, "The Effect of Modal Truncation on Modal Modification," *Proceedings of the 5th International Modal Analysis Conference*, SEM, pp. 72–78.

Ewins, D. J., ,1986, *Modal Testing: Theory and Practice*, John Wiley and Sons, New York.

Fried, I., 1979, *Numerical Solution of Differential Equations*, Academic Press, New York.

Gladwell, G. M. L., 1986, *Inverse Problems in Vibration*, Martinus Nijhoff Publishers, Dordrecht.

Golub, G. H., and Van Loan, C. F., 1983, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Md.

Ram, Y. M., Braun, S. G., and Blech, J. J., 1988, "Structural Modification in Truncated Systems by the Rayleigh-Ritz Method," *Journal of Sound and Vibration*, Vol. 125, No. 2, pp. 203–209.

Ram, Y. M., and Braun, S. G., 1989, "The Rayleigh-Ritz Method in Structural Modification," *Proceedings of the 7th International Modal Analysis Conference*, SEM, pp. 557–561.

Ram, Y. M., Blech, J. J., and Braun, S. G., 1990a, "Eigenproblem Error Bounds with Application to Symmetric Dynamic System Modification," *SIAM Journal on Matrix Analysis and Applications*, Vol. 11, No. 4, pp. 553–564.

Ram, Y. M., and Braun, S. G., 1990b, "Upper and Lower Bounds for the Natural Frequencies of Modified Structures Based on Truncated Modal Testing Results," *Journal of Sound and Vibration*, Vol. 137, No. 1, pp. 69–81.

Ram, Y. M., and Braun, S. G., 1990c, "Structural Dynamic Modification Using Truncated Data: Bounds for the Eigenvalues," *Mechanical Systems and Signal Processing*, Vol. 4, No. 1, pp. 39–52.

Zhang, Q., Wang, W., Allemang, R. J., and Brown, D. L., 1988, "Prediction of Mass Modification for Desired Natural Frequencies," Proceedings of the 6th International Modal Analysis Conference, pp. 1026–1032.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Subir Kumar Saha

Jorge Angeles
Mem. ASME

Department of Mechanical Engineering,
Robotic Mechanical Systems Laboratory,
McGill Research Centre for Intelligent Machines,
McGill University,
Montreal, Quebec, Canada

Dynamics of Nonholonomic Mechanical Systems Using a Natural Orthogonal Complement

The dynamics equations governing the motion of mechanical systems composed of rigid bodies coupled by holonomic and nonholonomic constraints are derived. The underlying method is based on a natural orthogonal complement *of the matrix associated with the velocity constraint equations written in linear homogeneous form. The method is applied to the classical example of a rolling disk and an application to a 2-dof Automatic Guided Vehicle is outlined.*

1 Introduction

The theory of nonholonomic systems arose when the analytical formalism of Euler and Lagrange was found to be inapplicable to the very simple mechanical problems of rigid bodies rolling without slipping on a plane. In fact, as late as 1894, Hertz (Neĭmark and Fufaev, 1967) introduced the distinction between holonomic and nonholonomic constraints in mechanical systems. Shortly thereafter, Čaplygin (1897) derived the dynamics equations in *true coordinates*, whereas Volterra (1898) derived the equations of motion in variables which he called *motion characteristics*. Appell (1899), on the other hand, proposed a new form of the equations of motion of nonholonomic systems while introducing the concept of *acceleration energy*, S, similar to kinetic energy, T. However, in spite of the simplicity of Appell's equations, it is harder to derive expressions for S than it is for T. A few years later, Maggi (1901) showed that Volterra's and Appell's equations may be derived from his method, first proposed in 1896. More recently, Kane (Kane, 1961; Kane and Wang, 1965) introduced a method for nonholonomic systems with elimination of constraint forces.

Neĭmark and Fufaev (1967) gave the first comprehensive and systematic exposition of the mechanics of nonholonomic systems, whereas Passerello and Huston (1973) expanded Kane's formulation by eliminating the computation of acceleration components. In their method, introduction of supplementary equations similar to the constraint relations with arbitrary choice of coefficients may be difficult; furthermore, the inversion of the associated matrix is unavoidable.

With the advent of digital computation, a series of new methods in the study of mechanical systems have been developed. Huston and Passerello (1974) introduced first a computer-oriented method similar to the method of the orthogonal complement of the matrix associated with the constraint equations, which reduces the dimension of the dynamical equations by elimination of constraint forces. Later, several formulations of dynamic modeling of closed-loop mechanical systems have been reported (Paul, 1975; Wehage and Haug, 1982; Kamman and Huston, 1984; Wampler et al., 1985; Kim and Vanderploeg, 1986). Each of those formulations are applicable to holonomic and nonholonomic systems with relative advantages and disadvantages.

On the other hand, the increasing need of dynamic simulation and control of robotic mechanical systems calls for efficient computational algorithms in this respect. As a matter of fact, current research interest in robotic mechanical systems with rolling contact, such as automated guided vehicles (AGV), has renewed the interest for the modeling and simulation of nonholonomic mechanical systems (Agulló et al., 1987, 1989; Muir and Neuman, 1987, 1988).

In this paper, a new method, based on *a natural orthogonal complement* (Angeles and Lee, 1988), which has already been applied to holonomic systems (Angles and Ma, 1988; Angeles and Lee, 1989), is applied to nonholonomic systems. The idea of the orthogonal complement of velocity constraints in the derivation of dynamical equations is not new, for it has been extensively used in multibody dynamics (Huston and Passerello, 1974; Hemami and Weimer, 1981; Kamman and Huston, 1984). Orthogonal complement-based methods of dynamics analysis consist of determining a matrix—an orthogonal complement—whose columns span the nullspace of the matrix of velocity constraints. However, the said orthogonal complement is not unique. In some approaches, an orthogonal complement is found with numerical schemes which are of an intensive nature, requiring, for example, singular-value decomposition or eigenvalue computations (Wehage and Haug, 1982; Kamman and Huston, 1984). In the recent approach, the orthogonal complement comes out naturally without any complex computations. The computation of both the *natural orthogonal complement* that we use and its time derivative is outlined in Section 4. The method is illustrated with the clas-

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 27, 1989; final revision, Nov. 17, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

sical problem of the rolling disk. As well, an application to a three-wheeled 2-dof AGV is outlined.

2 Introduction of a Natural Orthogonal Complement

In this paper, as a rule, we denote vectors with boldface lower cases, while tensors and matrices with boldface upper cases, regardless of the dimensions of the vectors and matrices involved. Explicit indication of these dimensions are mentioned when defined.

As pertaining to mechanical systems composed of constrained rigid bodies, the method of analysis based on the concept of *a natural orthogonal complement*, first introduced in (Angeles and Lee, 1988), is described briefly in the following steps:

Step 1: The *twist* of the ith rigid body of the system under study, undergoing an arbitrary motion in the three-dimensional space, $\mathbf{t}_i$, is defined in terms of its angular velocity, ω_i, and the velocity of the corresponding mass center, $\dot{\mathbf{c}}_i$, both being, in general, three-dimensional vectors. Hence, $\mathbf{t}_i$ is the following six-dimensional vector:

$$\mathbf{t}_i \equiv \begin{bmatrix} \omega_i \\ \dot{\mathbf{c}}_i \end{bmatrix}. \tag{1}$$

Moreover, if $\mathbf{I}_i$ denotes the 3×3 *inertia tensor* of the ith rigid body about its mass center, and this, as well as all vector quantities involved, are referred to a coordinate system fixed to the body, then, the Newton-Euler equations governing the motion of the ith body are written as follows:

$$\mathbf{M}_i \dot{\mathbf{t}}_i = -\mathbf{W}_i \mathbf{M}_i \mathbf{t}_i + \mathbf{w}_i \tag{2}$$

where the six-dimensional *wrench* vector, $\mathbf{w}_i$, acting on the ith body is defined, in accordance with the definition of $\mathbf{t}_i$, as

$$\mathbf{w}_i \equiv \begin{bmatrix} \mathbf{n}_i \\ \mathbf{f}_i \end{bmatrix}, \tag{3}$$

$\mathbf{n}_i$ and $\mathbf{f}_i$ being three-dimensional vectors, the former denoting the resultant moment, the latter the resultant force acting at the mass center of the ith body. Now, the 3×3 Cartesian tensor $\mathbf{\Omega}_i$ is defined as

$$\mathbf{\Omega}_i \equiv \frac{\partial(\omega_i \times \mathbf{x})}{\partial \mathbf{x}} \equiv \omega_i \times \mathbf{1} \tag{4}$$

for an arbitrary three-dimensional vector $\mathbf{x}$, whereas the 6×6 matrices of *extended angular velocity*, $\mathbf{W}_i$, and of *extended mass*, $\mathbf{M}_i$, are then defined as

$$\mathbf{W}_i \equiv \begin{bmatrix} \mathbf{\Omega}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{M}_i \equiv \begin{bmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & m_i \mathbf{1} \end{bmatrix}, \tag{5}$$

m_i, denoting the mass of the ith rigid body, whereas $\mathbf{0}$ and $\mathbf{1}$ denote the zero and the identity 3×3 tensors, respectively.

Step 2: If it is assumed that the mechanical system under study is composed of p rigid bodies, then the Newton-Euler equations for all individual bodies can be written as

$$\mathbf{M}_i \dot{\mathbf{t}}_i = -\mathbf{W}_i \mathbf{M}_i \mathbf{t}_i + \mathbf{w}_i^W + \mathbf{w}_i^N, \quad i = 1, \ldots, p \tag{6}$$

where $\mathbf{w}_i^W$ and $\mathbf{w}_i^N$ are the working wrench and the nonworking constraint wrench, both acting on the ith body, respectively. The former are understood as working moments and forces supplied by actuators or arising from dissipation; the latter, as nonworking moments and forces whose sole role is that of keeping the bodies together. Next, the $6p \times 6p$ matrices of *generalized mass*, $\mathbf{M}$, and of *generalized angular velocity*, $\mathbf{W}$, as well as the $6p$-dimensional vectors of *generalized twist*, $\mathbf{t}$, of *generalized working wrench*, $\mathbf{w}^W$, and *generalized nonworking constraint wrench*, $\mathbf{w}^N$, are defined as

$$\mathbf{M} \equiv \text{diag}\,[\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_p] \tag{7}$$

$$\mathbf{W} \equiv \text{diag}\,[\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_p] \tag{8}$$

$$\mathbf{t} \equiv \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_p \end{bmatrix}, \quad \mathbf{w}^W \equiv \begin{bmatrix} \mathbf{w}_1^W \\ \mathbf{w}_2^W \\ \vdots \\ \mathbf{w}_p^W \end{bmatrix}, \quad \mathbf{w}^N \equiv \begin{bmatrix} \mathbf{w}_1^N \\ \mathbf{w}_2^N \\ \vdots \\ \mathbf{w}_p^N \end{bmatrix}. \tag{9}$$

Hence, the p dynamical equations (6) can now be expressed in compact form as follows:

$$\mathbf{M}\dot{\mathbf{t}} = -\mathbf{W}\mathbf{M}\mathbf{t} + \mathbf{w}^W + \mathbf{w}^N \tag{10}$$

which is an equation formally identical to equation (2), and constitutes a set of $6p$ *unconstrained scalar* dynamical equations.

Step 3: The kinematic constraints produced by holonomic and nonholonomic couplings are derived in differential form. Within the methodology adopted here—as shown in Section 3—, every holonomic constraint gives rise to six scalar equations. As well, every nonholonomic constraint in the absence of slippage gives rise to three scalar equations. Moreover, due to the presence of the holonomic constraints, the overall constraint equations are not independent, and can be represented as a system of linear homogeneous equations on the twists. This is equivalent to the following linear homogeneous system on the vector of *generalized twist*:

$$\mathbf{A}\mathbf{t} = \mathbf{0}. \tag{11}$$

Here, $\mathbf{A}$ is a $(6\gamma + 3\nu) \times 6p$ matrix, γ and ν being the numbers of holonomic and nonholonomic couplings, respectively. Note that, with the approach introduced here, no distinction need be made between schleronomic and rheonomic constraints, for all are treated as schleronomic ones.

Step 4: Under the assumption that the degree-of-freedom of the system is n, an n-dimensional vector $\dot{\boldsymbol{\theta}}$ of *independent generalized speeds* is defined. Then, the vector of generalized twist can be represented as the following linear transformation of $\dot{\boldsymbol{\theta}}$:

$$\mathbf{t} = \mathbf{T}\dot{\boldsymbol{\theta}} \tag{12}$$

where $\mathbf{T}$ is a $6p \times n$ matrix. Upon substitution of $\mathbf{t}$, as given by equation (12), into equation (11), and recalling that all components of $\dot{\boldsymbol{\theta}}$ are independent, the following relation is readily derived:

$$\mathbf{A}\mathbf{T} = \mathbf{0} \tag{13}$$

which shows that $\mathbf{T}$ is an orthogonal complement of $\mathbf{A}$. Because of the particular form of choosing this complement—equation (12)—, $\mathbf{T}$ is termed *a natural orthogonal complement* of $\mathbf{A}$.

Step 5: By virtue of the definition of $\mathbf{A}$ and the vector of nonworking constraint wrench, the latter turns out to lie in the range of the transpose of $\mathbf{A}$ and hence, the said wrench lies in the nullspace of the transpose of $\mathbf{T}$. Therefore, upon multiplication of both sides of the $6p$-dimensional Newton-Euler uncoupled equations of the system, equation (10), by the transpose of $\mathbf{T}$, the vector of nonworking constraint wrench is eliminated from the said equation, which reduces to:

$$\mathbf{T}^T\mathbf{M}\dot{\mathbf{t}} = -\mathbf{T}^T\mathbf{W}\mathbf{M}\mathbf{t} + \mathbf{T}^T\mathbf{w}^W. \tag{14}$$

Step 6: Now, both sides of equation (12) are differentiated with respect to time, which yields

$$\dot{\mathbf{t}} = \mathbf{T}\ddot{\boldsymbol{\theta}} + \dot{\mathbf{T}}\dot{\boldsymbol{\theta}}. \tag{15}$$

Note that the elements of $\dot{\mathbf{T}}$ are not, in general, simply the time derivatives of the corresponding elements of $\mathbf{T}$, because the vector bases on which $\mathbf{T}$ is expressed are usually time varying.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

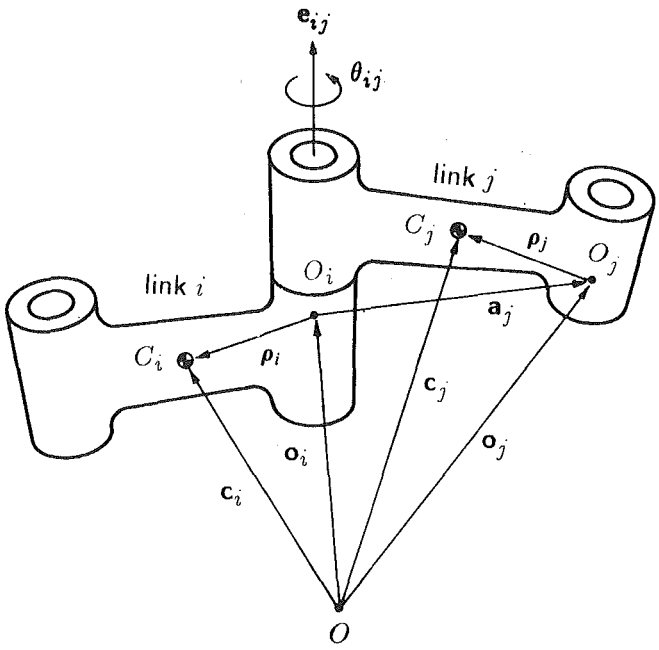


Fig. 1 Two links coupled by a revolute joint

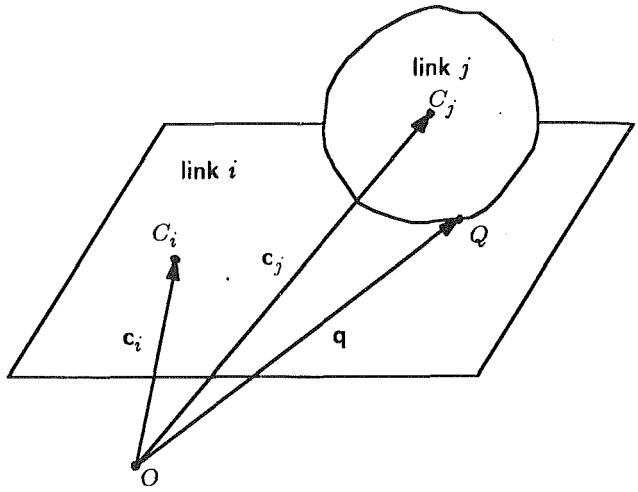


Fig. 2 A rigid body rolling on a plane

Furthermore, $\mathbf{w}^W$ is decomposed as follows:

$$\mathbf{w}^W = \mathbf{w}^A + \mathbf{w}^G + \mathbf{w}^D \tag{16}$$

where $\mathbf{w}^A$ represents the generalized wrench due to torques and forces applied by the actuators, if any, whereas $\mathbf{w}^G$ and $\mathbf{w}^D$ account for gravity and dissipative effects, respectively.

Upon substitution of equations (15) and (16) into equation (14), the following system of n-independent constrained dynamical equations are derived:

$$\mathbf{T}^T\mathbf{M}\mathbf{T}\ddot{\boldsymbol{\theta}} = -\mathbf{T}^T(\mathbf{M}\dot{\mathbf{T}} + \mathbf{W}\mathbf{M}\mathbf{T})\dot{\boldsymbol{\theta}} + \mathbf{T}^T(\mathbf{w}^A + \mathbf{w}^G + \mathbf{w}^D) \tag{17}$$

or,

$$\mathbf{I}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} = \mathbf{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})\dot{\boldsymbol{\theta}} + \boldsymbol{\tau} + \boldsymbol{\gamma} + \boldsymbol{\delta} \tag{18}$$

where

$\mathbf{I} \equiv \mathbf{T}^T\mathbf{M}\mathbf{T}$ $\equiv$ $n \times n$ matrix of generalized inertia.

$\mathbf{C} \equiv -\mathbf{T}^T(\mathbf{M}\dot{\mathbf{T}} + \mathbf{W}\mathbf{M}\mathbf{T})$ $\equiv$ $n \times n$ matrix of convective inertia terms.

$\boldsymbol{\tau} \equiv \mathbf{T}^T\mathbf{w}^A$ $\equiv$ n-dimensional vector of generalized driving force.

$\boldsymbol{\gamma} \equiv \mathbf{T}^T\mathbf{w}^G$ $\equiv$ n-dimensional vector of generalized force due to gravity.

$\boldsymbol{\delta} \equiv \mathbf{T}^T\mathbf{w}^D$ $\equiv$ n-dimensional vector of generalized dissipative force.

From the foregoing discussion, then, it becomes apparent that equation (18) represents the system's Euler-Lagrange dynamical equations, free of nonworking generalized constraint forces.

3 Derivation of the Kinematic Constraints

Mechanical couplings that produce holonomic and nonholonomic constraints on the twists—velocity constraints—of the coupled bodies, say $\mathbf{t}_i$ and $\mathbf{t}_j$ defined as in equation (1), have the following form:

$$\mathbf{A}_i\mathbf{t}_i + \mathbf{A}_j\mathbf{t}_j = 0. \tag{19}$$

The foregoing constraint equations are seen to be *linear homogeneous* in the twists of the coupled bodies. Moreover, the coefficient matrices $\mathbf{A}_i$ and $\mathbf{A}_j$ are, in general, configuration dependent.

3.1 Holonomic Systems. In the case of a holonomic coupling, for example, a revolute pair, the constraint equations are readily derived as follows: If $\dot{\theta}_{ij}$ is the joint rate for the revolute coupling between the ith and the jth links, then, referring to Fig. 1, the relative angular velocity of the jth link with respect to the ith link, $\boldsymbol{\omega}_j - \boldsymbol{\omega}_i$, is $\dot{\theta}_{ij}\mathbf{e}_{ij}$. Thus, the equation constraining the angular velocities of two successive links is the following:

$$\mathbf{e}_{ij} \times (\boldsymbol{\omega}_j - \boldsymbol{\omega}_i) = 0. \tag{20}$$

Furthermore, from Fig. 1 it is clear that

$$\dot{\mathbf{c}}_j = \dot{\mathbf{c}}_i + \boldsymbol{\omega}_j \times (\mathbf{a}_j + \boldsymbol{\rho}_j) - \boldsymbol{\omega}_i \times \boldsymbol{\rho}_i. \tag{21}$$

Equations (20) and (21) are now written in terms of the link twists, which readily produces an equation of the form of equation (19), where the 6×6 matrices $\mathbf{A}_i$ and $\mathbf{A}_j$ are as shown below:

$$\mathbf{A}_i = \begin{bmatrix} -\mathbf{e}_{ij} \times \mathbf{1} & \mathbf{0} \\ -\boldsymbol{\rho}_i \times \mathbf{1} & -\mathbf{1} \end{bmatrix}, \quad \mathbf{A}_j = \begin{bmatrix} \mathbf{e}_{ij} \times \mathbf{1} & \mathbf{0} \\ (\mathbf{a}_j + \boldsymbol{\rho}_j) \times \mathbf{1} & \mathbf{1} \end{bmatrix}. \tag{22}$$

Mechanical couplings that produce holonomic constraints other than the revolute pair are, e.g., the prismatic pair and gear trains. The pulley-belt and the cam-follower transmissions produce holonomic constraints as well. The different forms of the 6×6 matrices $\mathbf{A}_i$ and $\mathbf{A}_j$ for different couplings can be obtained in a similar way, as described in Angeles and Lee (1989).

3.2 Nonholonomic Systems. We will limit the discussion to nonholonomic pure rolling, and hence, no slippage is considered. As an example of this type of nonholonomic coupling, a rigid body rolling without slipping on a plane is shown in Fig. 2, the nonslip condition being derived below. If the rigid body is considered as the jth body of a system and the plane is a part of the boundary of the ith body, then, referring to Fig. 2, the arising nonholonomic constraint is stated as

$$\dot{\mathbf{c}}_j = \dot{\mathbf{c}}_i + \boldsymbol{\omega}_j \times (\mathbf{c}_j - \mathbf{q}) - \boldsymbol{\omega}_i \times (\mathbf{c}_i - \mathbf{q}) \tag{23}$$

where $\mathbf{q}$ is the position vector of the contact point Q. Moreover, equation (23) can also be written as equation (19), with the 3×6 matrices $\mathbf{A}_i$ and $\mathbf{A}_j$ defined as:

$$\mathbf{A}_i = [-\mathbf{C}_i \ -\mathbf{1}], \quad \mathbf{A}_j = [\mathbf{C}_j \ \mathbf{1}] \tag{24}$$

and the Cartesian tensors $\mathbf{C}_i$ and $\mathbf{C}_j$ defined as follows:

$$\mathbf{C}_i = (\mathbf{c}_i - \mathbf{q}) \times \mathbf{1}, \quad \text{and} \quad \mathbf{C}_j = (\mathbf{c}_j - \mathbf{q}) \times \mathbf{1}. \tag{25}$$

From equations (22) and (24) it is apparent that a kinematic constraint, whether holonomic or nonholonomic, can be writ-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ten in the form of a linear homogeneous equation involving the twists of every pair of coupled bodies of a mechanical system. An essential difference between the holonomic and the nonholonomic constraints previously derived is pointed out here. Whereas the former lead to six-dimensional constraint equations, the latter lead to three-dimensional equations. As one can readily verify, the holonomic constraints derived above are not independent, although the nonholonomic constraints are. If slippage is possible, then the number of independent nonholonomic constraint equations diminishes, for the relative motion gains degrees-of-freedom. Furthermore, if the kinematic constraints introduced by the couplings comprise γ holonomic and ν nonholonomic couplings, then the arising kinematic constraint equations can be described in compact form as in equation (11), where $\mathbf{A}$ is a $(6\gamma + 3\nu) \times 6p$ matrix. This approach departs from that reported in the literature of deriving *independent* scalar constraint equations that are linear in the body twists, but *not necessarily homogeneous*. In fact, the homogeneity of these constraint equations allows us to regard all constraints as schleronomic.

4 Calculation of T and $\dot{\mathbf{T}}$

The derivation of the *natural orthogonal complement* from equation (12) is, in general, costly, except for simple systems, as in the example given below. An efficient method of calculating $\mathbf{T}$ numerically (Ma and Angeles, 1989) can be readily derived by noticing that $\mathbf{T}$ depends on *generalized coordinates* only. Moreover, the jth column of $\mathbf{T}$ equals $\partial \mathbf{t}/\partial \dot{\theta}_j$, for $j = 1, 2, \ldots, n$. Thus, $\mathbf{T}$ can be found as follows:

$$\mathbf{T} = [\mathbf{t}|_{\dot{\theta}_1=1},\ \mathbf{t}|_{\dot{\theta}_2=1},\ \ldots,\ \mathbf{t}|_{\dot{\theta}_n=1}]_{\text{other } \dot{\theta}\text{'s of } \dot{\theta} \text{ are zero}}, \quad (26)$$

i.e., the jth column of $\mathbf{T}$ is calculated as the generalized twist of the system assuming that all the independent speeds are zero but the jth one, which has a value of unity.

In simulation applications, there is no need of an explicit evaluation of $\dot{\mathbf{T}}$. In fact, the generalized inertia terms that are quadratic in $\dot{\theta}$, $\mathbf{C}(\theta, \dot{\theta})\dot{\theta}$, of equation (18), can be evaluated efficiently with the technique introduced by Walker and Orin (1982) for serial manipulators and which has been extended by Ma and Angeles (1989) for parallel manipulators. This is given as follows:

$$\mathbf{C}(\theta, \dot{\theta})\dot{\theta} = \tau + \gamma + \delta|_{\ddot{\theta}=0}, \quad (27)$$

i.e., the entire first term of equation (18) is computed from *inverse dynamics* (Walker and Orin, 1982) as the force, consisting of driving, gravity, and dissipative force, if any, respectively, required to maintain the system's current joint displacements and velocities, but with zero accelerations $\ddot{\theta}$.

5 Example: Disk Rolling on an Inertial Plane

Shown in Fig. 3 is a disk rolling on a plane that is considered fixed to an inertial frame. Then, the mass center of the disk, C, is assumed to be coincident with its centroid, the velocity of the latter being denoted by $\dot{\mathbf{c}}$. Moreover, vector $\dot{\mathbf{c}}$ is related to the disk's angular velocity, ω, as follows:

$$\dot{\mathbf{c}} = \omega \times (\mathbf{c} - \mathbf{q}) \quad (28)$$

where $\mathbf{c}$ and $\mathbf{q}$ are position vectors of the mass center C and the point of contact Q, respectively. Since p, the number of moving rigid bodies, is equal to one, i.e., the disk, the six-dimensional vector of twist, $\mathbf{t}$, similar to equation (12), can be written as a linear transformation of the *independent generalized speeds*. We can choose, for example, the angular velocity ω of the disk as the vector of independent generalized speeds, because the degree-of-freedom of the system is three. Now, $\mathbf{t}$ is given by

$$\mathbf{t} = \mathbf{T}\omega \quad (29)$$

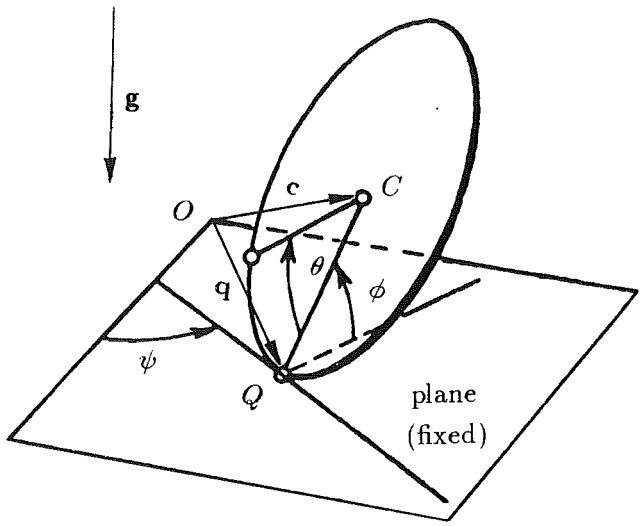


Fig. 3 A disk rolling on an inertial plane

with the 6×3 matrix $\mathbf{T}$ defined as:

$$\mathbf{T} = \begin{bmatrix} \mathbf{1} \\ (\mathbf{q}-\mathbf{c}) \times \mathbf{1} \end{bmatrix}. \quad (30)$$

Next, the constraint equation is simply derived as:

$$\mathbf{A}\mathbf{t} = \mathbf{0} \quad (31)$$

where

$$\mathbf{A} = [(\mathbf{c}-\mathbf{q}) \times \mathbf{1}\ \mathbf{1}]$$

and hence, $\mathbf{A}$ is a 3×6 matrix. Matrix $\mathbf{T}$ is an orthogonal complement of $\mathbf{A}$, which can easily be proven by simply performing the product $\mathbf{AT}$.

Now, the 6×6 *generalized mass and generalized angular velocity* matrices $\mathbf{M}$ and $\mathbf{W}$ are defined as follows:

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_C & \mathbf{0} \\ \mathbf{0} & m\mathbf{1} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \omega \times \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where m and $\mathbf{I}_C$ are the mass and the 3×3 inertia tensor about the mass center of the disk, respectively. To obtain the equations of motion, $\dot{\mathbf{T}}$ is calculated as:

$$\dot{\mathbf{T}} = \begin{bmatrix} \mathbf{0} \\ -\dot{\mathbf{c}} \times \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -[\omega \times (\mathbf{c}-\mathbf{q})] \times \mathbf{1} \end{bmatrix}.$$

Then, the 3×3 *generalized inertia* matrix, $\mathbf{I}$, as given in equation (18), is derived as:

$$\mathbf{I} = [\mathbf{1} - (\mathbf{q}-\mathbf{c}) \times \mathbf{1}] \begin{bmatrix} \mathbf{I}_C & \mathbf{0} \\ \mathbf{0} & m\mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ (\mathbf{q}-\mathbf{c}) \times \mathbf{1} \end{bmatrix}$$

$$= \mathbf{I}_C + m[(\mathbf{q}-\mathbf{c}) \cdot (\mathbf{q}-\mathbf{c})\mathbf{1} - (\mathbf{q}-\mathbf{c}) \otimes (\mathbf{q}-\mathbf{c})] \quad (32)$$

where $\otimes$ denotes the *tensor product* of the two vectors beside it. Equation (32) has the following interpretation: The calculated matrix of *generalized inertia* $\mathbf{I}$ is nothing but the moment of inertia of the disk about the contact point Q. Moreover, the matrix $\mathbf{C}$ of convective inertia terms defined in equation (18), is derived as follows:

$$\mathbf{WMT} = \begin{bmatrix} \omega \times \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I}_C & \mathbf{0} \\ \mathbf{0} & m\mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ (\mathbf{q}-\mathbf{c}) \times \mathbf{1} \end{bmatrix} = \begin{bmatrix} \omega \times \mathbf{I}_C \\ \mathbf{0} \end{bmatrix}$$

and $\mathbf{M}\dot{\mathbf{T}} =$

$$\begin{bmatrix} \mathbf{I}_C & \mathbf{0} \\ \mathbf{0} & m\mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ -[\omega \times (\mathbf{q}-\mathbf{c})] \times \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -m[\omega \times (\mathbf{q}-\mathbf{c})] \times \mathbf{1} \end{bmatrix}.$$

Therefore,

$$\mathbf{C} = -\mathbf{T}^T(\mathbf{WMT} + \mathbf{M}\dot{\mathbf{T}})$$

$$= -\omega \times \mathbf{I}_C - m[(\mathbf{q}-\mathbf{c}) \times \omega] \otimes (\mathbf{q}-\mathbf{c}). \quad (33)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

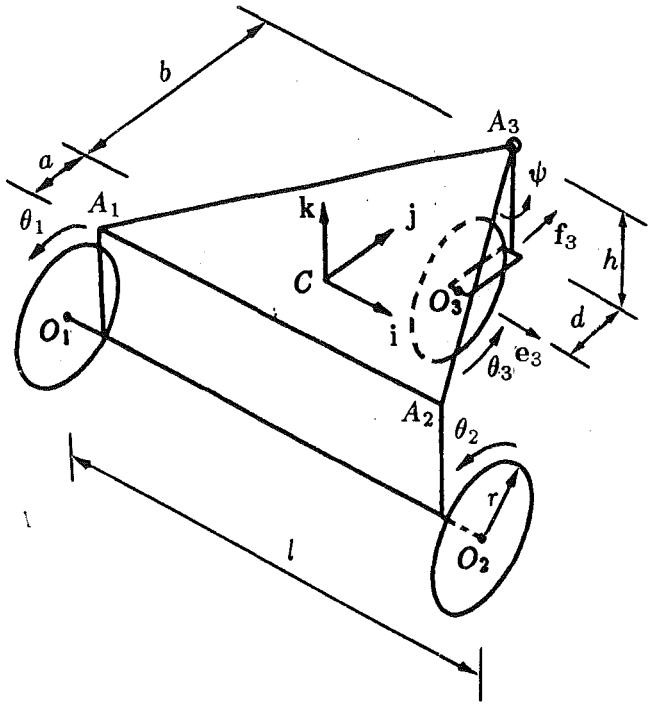


Fig. 4 Three-dimensional view of a 2-dof AGV

If dissipation is neglected and there is no actuation, the three-dimensional vectors δ and τ vanish but γ, the *generalized force* due to gravity, is given as:

$$\gamma = \mathbf{T}^T \mathbf{w}^W = [\mathbf{1} - (\mathbf{q} - \mathbf{c}) \times \mathbf{1}] \begin{bmatrix} \mathbf{0} \\ m\mathbf{g} \end{bmatrix} = m(\mathbf{c} - \mathbf{q}) \times \mathbf{g} \quad (34)$$

where $\mathbf{g}$ is the vector of gravity acceleration. Thus, γ is nothing but the moment of the gravity force about the contact point Q.

Finally, the equation of motion of the disk rolling on a plane without slipping can be written as:

$$\mathbf{I}\omega = \mathbf{C}\omega + \gamma \quad (35)$$

where $\mathbf{I}$, $\mathbf{C}$, and γ are as derived in equations (32), (33) and (34), respectively.

6 Application to a 2-DOF Automatic Guided Vehicle: An Outline

The kinematics and dynamics analyses of a three-wheeled 2-dof AGV, using the natural orthogonal complement introduced above, have been discussed in detail in (Saha and Angeles, 1989). An outline of the modeling of the same vehicle is included here for illustration purposes.

Step 1: It is assumed that the AGV under study, as shown in Fig. 4, contains a platform, two rear driving wheels, and a caster wheel in the front. Here, note that not all joint motions are actuated or driven by motors. Only the two rear wheels are actuated by independent motors whose motions are grouped in the two-dimensional vector of actuated joint motions, $\theta_a = [\theta_1, \theta_2]^T$, in agreement with the fact that the vehicle has two degrees-of-freedom. The other two joint motions, spinning of the front wheel, θ_3, and rotation of the bracket holding the front caster wheel with respect to the platform, ψ, are unactuated joint motions. Thus, the angular displacements of the actuated joints can be considered as the independent joint variables and their time rates as the independent speeds. Now, considering the motion of the vehicle being planar, the three-dimensional *reduced twist* vector, $\mathbf{t'}_p = [\omega, v_i, v_j]^T$, is defined, where ω denotes the scalar angular velocity of the platform, while v_i and v_j are the velocity components of mass-center of the platform in the directions of $\mathbf{i}$ and $\mathbf{j}$, respectively. Vector $\mathbf{t'}_p$ is now written as a linear transformation of the two-dimensional vector of actuated joint rates, $\dot{\theta}_a$. This relation, which was derived in (Saha and Angeles, 1989), is given as

$$\mathbf{t'}_p = \mathbf{T'}_p \dot{\theta}_a \quad (36)$$

where $\mathbf{T'}_p$ is a 3×2 matrix.

Step 2: For inverse kinematics, $\dot{\theta}_a$ is to be calculated for given $\mathbf{t}_p'$. Now, for kinematically admissible motions, it can be proven that $\mathbf{t}_p'$ lies in the range of $\mathbf{T}_p'$. Therefore, the three components of $\mathbf{t'}_p$ cannot be supplied arbitrarily, which is evidenced by the fact that the AGV has only two degrees-of-freedom. Hence, the *actuated joint-rate* vector, $\dot{\theta}_a$, can be calculated as the least-square *approximation* of the overdetermined system of equations, appearing in equation (36), with zero error. This means that the least-square approximation of that system is, in fact, its solution, for this system is overdetermined only formally.

Step 3: The relation in equation (36) is differentiated with respect time. This is clearly as follows:

$$\dot{\mathbf{t}}'_p = \mathbf{T'}_p \ddot{\theta}_a + \dot{\mathbf{T}}'_p \dot{\theta}_a. \quad (37)$$

Then, once the solution for $\dot{\theta}_a$ is known from Step 2, $\ddot{\theta}_a$ can be solved from the equation (37) following exactly the same least-square approach of Step 2.

Step 4: To obtain the two-dimensional *actuated joint-angle* vector, θ_a, the expression for $\dot{\theta}_a$ obtained from equation (36) is integrated, with known initial conditions, by any standard integration scheme.

Step 5: The dynamical equations of motion, which are derived in Saha and Angeles (1989), can now be written as:

$$\mathbf{I}(\theta)\ddot{\theta}_a = \mathbf{C}(\theta, \dot{\theta}_a)\dot{\theta}_a + \tau_a \quad (38)$$

where $\mathbf{I}$ and $\mathbf{C}$ are 2×2 matrices of *generalized inertia* and of *convective inertia terms*, respectively. Moreover, τ_a is a two-dimensional vector of *generalized torque* supplied by the actuators, θ is the four-dimensional vector of actuated and unactuated joint angles, whereas $\dot{\theta}_a$ and $\ddot{\theta}_a$ are the two-dimensional vectors of actuated joint rates and accelerations, respectively.

Step 6: With the above mentioned methodology, the simulation of the automatic guided vehicle under study, when traversing a circular path (Saha and Angeles, 1989) and two parallel straight lines connected with a smooth curve, has been performed, the results not being included here due to space limitations.

7 Conclusions

In this paper, the derived dynamical model of mechanical systems with nonholonomic and holonomic constraints, equation (18), using *a natural orthogonal complement*, were derived free of constraint forces, which usually appear when the Newton-Euler method is applied to derive the governing equations. Moreover, what the derivation of the dynamical equations presented here shows is that these can be obtained without resorting to lengthy partial differentiations, which would be the case if either a straightforward or a recursive derivation of the Euler-Lagrange equations were attempted. From an algorithmic point of view, the method of the *natural orthogonal complement* has advantages over other, similar methods, as pointed out in Section 4. Here, $\mathbf{T}$ is calculated efficiently by avoiding eigenvalue or singular-value calculations. Moreover, in simulation applications the convective inertia terms, which involve the evaluation of $\dot{\mathbf{T}}$, as well as a few more matrix-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

times-vector multiplications and vector additions, is calculated very efficiently using Walker and Orin's scheme (Walker and Orin, 1982). On the other hand, the direct application of Kane's method (Kane, 1961; Kane and Wang, 1965) requires the evaluation of acceleration terms, where the method of Passerello and Huston (1973) requires a matrix inversion. None of these is needed when *a natural orthogonal complement* is used.

Acknowledgments

The research work reported here was possible under NSERC Grant No. A4532 and FCAR (Fonds pour la formation de chercheurs et l'aide à la recherche, of Québec) Grant No. 88AS2517. The scholarship granted to Subir Kumar Saha by the Government of India is highly acknowledged.

References

Agulló, J., Cardona, S., and Vivancos, J., 1987, "Kinematics of vehicles with directional sliding wheels," *Mechanism and Machine Theory*, Vol. 22, No. 4, pp. 295–301.

Agulló, J., Cardona, S., and Vivancos, J., 1989, "Dynamics of vehicles with directionally sliding wheels," *Mechanism and Machine Theory*, Vol. 24, No. 1, pp. 53–60.

Angles, J., and Lee, S., 1988a, "The formulation of dynamical equations of holonomic mechanical systems using a natural orthogonal complement," ASME Journal of Applied Mechanics, Vol. 55, pp. 243–244.

Angeles, J. and Lee, S., 1989, "The modelling of holonomic mechanical systems using a natural orthogonal complement," *Trans. Canadian Society of Mechanical Engineers*, Vol. 13, No. 4, pp. 81–89.

Angeles, J., and Ma, O., 1988, "Dynamic simulation of n-axis serial robotic manipulators using a natural orthogonal complement," *The Int. J. of Robotics Research*, Vol. 7, No. 5, pp. 32–47.

Appell, P., 1899, "Sur une forme nouvelle des équations de la dynamique," *J. Reine Angew. Math.*, Vol. 121, pp. 459–460.

Čaplygin, S. A., 1897, "On the motion of a heavy figure of revolution on a horizontal plane" (in Russian), *Trudy Otd. Fiz. Nauk Obšč. Ljubitel. Estest*, Vol. 9, pp. 10–16.

Hemami, H., and Weimer, F. C., 1981, "Modeling of nonholonomic dynamic systems with applications," ASME Journal of Applied Mechanics, Vol. 48, pp. 177–182.

Huston, R. L., and Passerello, C. E., 1974, "On Constraint Equations—A New Approach," ASME Journal of Applied Mechanics, Vol. 41, pp. 1130–1131.

Kamman, J. W., and Huston, R. L., 1984, "Dynamics of constrained multibody systems," ASME Journal of Applied Mechanics, Vol. 51, pp. 899–903.

Kane, T. R., 1961, "Dynamics of nonholonomic systems," Vol. 83, ASME Journal of Applied Mechanics, Vol. 28, pp. 574–578.

Kane, T. R., and Wang, C. F., 1965, "On the derivation of equations of motion," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 13, No. 2, pp. 487–492.

Kim, S. S., and Vanderploeg, M. J., 1986, "A general and efficient method for dynamic analysis of mechanical systems using velocity transformations," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 108, pp. 176–182.

Ma, O., and Angeles, J., 1989, "Direct kinematics and dynamics of a planar three-dof parallel manipulator," *Advances in Design Automation, Proc. of the 15th Design Automation Conference*, Montreal, Sept., pp. 313–320.

Maggi, G. A., 1901, "Di alcune nuove forme delle equazioni della dinamica applicabili ai sistemi anolonomi," *Atti Accad. Naz. Lincei Rend Cl. Fis. Mat. Nat.*, Vol. 10, No. 5, pp. 287–291.

Muir, P. F., and Neuman, C. P., 1987, "Kinematic modeling of wheeled mobile robots," *Journal of Robotic Systems*, Vol. 4, No. 2, pp. 281–340.

Muir, P. F., and Neuman, C. P., 1988, "Dynamic modeling of multibody roboticf mechanisms: incorporating closed-chains, friction, higher-pair joints, and unactuated and unsensed joints," *Proc. of the IEEE Conf. on Robotics and Automation*, Philadelphia, Pa., Apr. 24–29, pp. 1546–1551.

Neimark, Ju. I., and Fufaev, N. A., 1967, *Dynamics of Nonholonomic Systems* (in Russian), (Translated version in English, 1972), American Mathematical Society, Providence, R.I.

Passerello, C. E., and Huston, R. L., 1973, "Another look at nonholonomic systems," ASME Journal of Applied Mechanics, March, pp. 101–104.

Paul, B., 1975, "Analytical dynamics of mechanisms—A computer oriented overview," *Mechanism and Machine Theory*, Vol. 10, pp. 481–507.

Saha, S. K., and Angeles, J., 1989, "Kinematics and dynamics of 3-wheeled 2-DOF AGV," *Proc. of the IEEE Conf. on Robotics and Automation*, Scottsdale, Ariz., May, 14–19, pp. 1572–1577.

Volterra, V., 1898, "Sopra una classe di equazione dinamiche," *Atti Accad. Sci. Torino*, Vol. 33, pp. 451–475; errata, ibid, Vol. 35 (1899), p. 118.

Walker, M. W., and Orin, D. E., 1982, "Efficient dynamic computer simulation of robotic mechanisms," *ASME Journal of Dynamic Systems, Measurement, and Control*, Vol. 104, pp. 205–211.

Wampler, C., Buffinton, K., and Shu-hui, J., 1985, "Formulation of equations of motion for systems subject to constraints," ASME Journal of Applied Mechanics, Vol. 52, pp. 465–470.

Wehage, R. A., and Haug, E. J., 1982, "Generalized coordinate partitioning for dimension reduction in analysis of constrained dynamic systems," *ASME Journal of Mechanical Design*, Vol. 104, pp. 247–255.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Chaos in a Weakly Nonlinear Oscillator With Parametric and External Resonances

K. Yagasaki
Assistant Professor,
Department of Mechanical Engineering,
Tamagawa University,
Machida, Tokyo 194, Japan

This paper describes a study of the chaotic dynamics of a weakly nonlinear single degree-of-freedom system subjected to combined parametric and external excitation. We consider a case of double resonance in which primary resonances, with respect to parametric and external forces, exist simultaneously. By using the averaging method and Melnikov's technique, it is shown that chaos may occur in certain parameter regions. These chaotic motions result from the existence of orbits homoclinic to a normally hyperbolic invariant torus which corresponds to a hyperbolic periodic orbit in the averaged system. The mechanism and structure of chaos in this situation are also described. Furthermore, the existence of steady-state chaos is demonstrated by numerical simulation.

1 Introduction

In this paper we study the chaotic dynamics of a weakly nonlinear single degree-of-freedom system subjected to combined parametric and external excitation:

$$\ddot{x}+\bar{\delta}\dot{x}+(1+\bar{\beta}\cos\bar{\nu}t)x+\bar{\alpha}x^3=\bar{\gamma}\cos\omega t, \tag{1}$$

where an overdot denotes differentiation with respect to time t and damping $\bar{\delta}$, nonlinearity $\bar{\alpha}$, and parametric and external force amplitudes $\bar{\beta}$, $\bar{\gamma}$ are assumed to be small. In the previous study, Yagasaki et al. (1990) analyzed equation (1) for the case in which the primary resonance to external force exists and the parametric force frequency is low, i.e., $\omega\approx 1$ and $\bar{\nu}<<1$, and discussed quasi-periodic and chaotic motions. Here we consider a case of double resonance in which primary resonances to parametric and external forces exist simultaneously, i.e., $\bar{\nu}\approx 2$ and $\omega\approx 1$. By using the averaging method and a global perturbation technique originally due to Melnikov (1963), it is shown that chaotic motions may occur in certain parameter regions.

Since the frequencies ω and $\bar{\nu}$ are generally incommensurate, we must consider equation (1) as a nonlinear oscillator with quasi-periodic forcing. Ness (1971) and Troger and Hsu (1977) used the averaging method to obtain the steady-state solutions to equations of type (1). HaQuang et al. (1987) also analyzed systems similar to equation (1) using the method of multiple scales, and observed chaotic motions in numerical simulation. However, in these earlier works they did not take account of the quasi-periodicity of systems forced with two frequencies.

The chaotic dynamics of periodically forced systems has been extensively studied (cf., Guckenheimer and Holmes, 1983), but there has been less attention to quasi-periodically forced systems. Recently, Wiggins (1988a, 1988b), and Ide and Wiggins (1989) studied the chaotic dynamics of quasi-periodically forced systems with strong nonlinearity. They extended Melnikov's method to obtain a criterion for chaos to exist. Scheurle (1986) also examined chaotic orbits of an almost periodically forced oscillator utilizing a Melnikov-type technique. Their approach, however, cannot be directly applied to weakly nonlinear systems such as equation (1).

This paper is organized as follows: In Section 2 equation (1) is transformed into the form of an oscillator forced at high and low frequencies and the averaging method is applied to only fast oscillating terms. The averaged system takes the form of a time-periodic perturbation of an integrable Hamiltonian system with separatrix loops. In Section 3 we predict the regions in parameter space where the averaged system has transverse homoclinic orbits using the standard Melnikov method. The existence of such orbits implies that chaotic motions of Smale horseshoe type may occur in the averaged system. In Section 4 the chaotic dynamics of the weakly nonlinear oscillator (1) is described. These chaotic motions result from the existence of orbits homoclinic to a normally hyperbolic invariant torus which corresponds to a hyperbolic periodic orbit in the averaged system. The mechanism and structure of chaos in this situation are also described. In Section 5 the existence of steady-state chaos in the averaged system and the original system (1) is demonstrated by numerical simulation. We close a summary and some comments in Section 6.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Joint Applied Mechanics/Bioengineering Conference, Ohio State University, Columbus, Ohio, June 16-19, 1991.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 17, 1989; final revision, Mar. 1, 1990.

Paper No. 90-APM-3.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

2 Equation of Motion and Averaging

Introduce two small parameters ϵ and μ, such that $0<\epsilon<<\mu<<1$ and let $\bar{\alpha}=\epsilon\alpha$, $\bar{\gamma}=\epsilon\gamma$, $\bar{\delta}=\epsilon\mu\delta$, and $\bar{\beta}=\epsilon\mu\beta$. Consider simultaneous resonances in which both parametric and external resonances exist, i.e., $\bar{\nu}\approx 2$ and $\omega\approx 1$, and set $\bar{\nu}=2\omega+\epsilon\nu$ and $\omega^2-1=\epsilon\Omega$. Then equation (1) becomes

$$\ddot{x}+\epsilon\mu\delta\dot{x}+[1+\epsilon\mu\beta\cos(2\omega+\epsilon\nu)t]x+\epsilon\alpha x^3=\epsilon\gamma\cos\omega t,$$

or, as a first-order equation

$$\dot{x}=y, \tag{2}$$
$$\dot{y}=-x+\epsilon\{[-\alpha x^3+\gamma\cos\omega t]+\mu[-\delta y-\beta x\cos(2\omega t+\epsilon\nu t)]\}.$$

The system (2) contains a "fast" time t and a "slow" time ϵt.

We define a Poincaré map $P_{\epsilon,\mu}$ associated with (2). By extending the phase space to $\mathbf{R}^2\times S^1\times S^1$, the system (2) can be rewritten as

$$\dot{x}=y,$$
$$\dot{y}=-x+\epsilon\{[-\alpha x^3+\gamma\cos\theta]+\mu[-\delta y-\beta x\cos(2\theta+\tau)]\},$$
$$\dot{\tau}=\epsilon\nu,$$
$$\dot{\theta}=\omega,\ (x,y,\tau,\theta)\in\mathbf{R}^2\times S^1\times S^1, \tag{3}$$

where S^1 is the circle of length 2π and $\tau=\epsilon\nu t$, $\theta=\omega t \bmod 2\pi$. Define a global cross-section $\Sigma_\theta=\{(x,y,\tau,\theta)\in\mathbf{R}^2\times S^1\times S^1\,|\,\theta=0\}$. Then, the Poincaré map $P_{\epsilon,\mu}$: $\Sigma_\theta\to\Sigma_\theta$ is obtained by

$$P_{\epsilon,\mu}:\ (x(0),y(0),\tau)\to(x(T),y(T),\tau+\epsilon\nu T), \tag{4}$$

where $(x(t),\ y(t))$ is a solution of equation (2) and $T=2\pi/\omega$.

On the other hand, by using the invertible transformation

$$J=\frac{x^2+y^2}{2},\quad \phi=-\tan\left(\frac{y}{\omega x}\right)-\omega t, \tag{5a}$$

or

$$x=\sqrt{2J}\cos(\omega t+\phi),\quad y=-\omega\sqrt{2J}\sin(\omega t+\phi), \tag{5b}$$

equation (2) becomes

$$\dot{J}=-\frac{\epsilon\sqrt{2J}}{\omega}\{\Omega\sqrt{2J}\cos(\omega t+\phi)+\gamma\cos\omega t$$
$$-\alpha(2J)^{3/2}\cos^3(\omega t+\phi)+\mu[\delta\omega\sqrt{2J}\sin(\omega t+\phi)$$
$$-\beta\sqrt{2J}\cos(\omega t+\phi)\cos(2\omega+\epsilon\nu)t]\}\sin(\omega t+\phi),$$
$$\dot{\phi}=-\frac{\epsilon}{\omega\sqrt{2J}}\{\Omega\sqrt{2J}\cos(\omega t+\phi)+\gamma\cos\omega t$$
$$-\alpha(2J)^{3/2}\cos^3(\omega t+\phi)+\mu[\delta\omega\sqrt{2J}\sin(\omega t+\phi)$$
$$-\beta\sqrt{2J}\cos(\omega t+\phi)\cos(2\omega+\epsilon\nu)t]\}\cos(\omega t+\phi). \tag{6}$$

Applying the averaging procedure to only the fast time, one obtains

$$\dot{J}=\frac{\epsilon}{2\omega}\{-\gamma\sqrt{2J}\sin\phi-\mu[2\delta\omega J+\beta J\sin(\epsilon\nu t-2\phi)]\},$$
$$\dot{\phi}=\frac{\epsilon}{2\omega}\left\{-\Omega+\frac{3}{2}\alpha J-\frac{\gamma\cos\phi}{\sqrt{2J}}+\mu\frac{\beta}{2}\cos(\epsilon\nu t-2\phi)\right\}. \tag{7}$$

The averaging theorem (Hale, 1969; Guckenheimer and Holmes, 1983; Sanders and Verhulst, 1985) implies that there exists a near-identity transformation under which equation (6) becomes

$$\dot{J}=\frac{\epsilon}{2\omega}\{-\gamma\sqrt{2J}\sin\phi-\mu[2\delta\omega J+\beta J\sin(\epsilon\nu t-2\phi)]\}+0(\epsilon^2),$$
$$\dot{\phi}=\frac{\epsilon}{2\omega}\left\{-\Omega+\frac{3}{2}\alpha J-\frac{\gamma\cos\phi}{\sqrt{2J}}+\mu\frac{\beta}{2}\cos(\epsilon\nu t-2\phi)\right\}+0(\epsilon^2), \tag{8}$$

where the $0(\epsilon^2)$ remainder terms are quasi-periodic in t with two frequencies ω and $\epsilon\nu$, and that solutions of (6) are well approximated by solutions of (7) (see equation (26) in Section 4). The $0(\epsilon\mu)$ terms in equation (8) are retained, while the $0(\epsilon^2)$ terms neglected. This is justified by the scaling $0<\epsilon<<\mu<<1$.

By changing time variable $t\to\epsilon t/2\omega$, equation (7) becomes

$$\dot{J}=-\gamma\sqrt{2J}\sin\phi-\mu[2\delta\omega J+\beta J\sin(\nu_0 t-2\phi)],$$
$$\dot{\phi}=-\Omega+\frac{3}{2}\alpha J-\frac{\gamma\cos\phi}{\sqrt{2J}}+\mu\frac{\beta}{2}\cos(\nu_0 t-2\phi) \tag{9}$$

where $\nu_0=2\omega\nu$. For $\mu=0$, one obtains

$$\dot{J}=-\gamma\sqrt{2J}\sin\phi,$$
$$\dot{\phi}=-\Omega+\frac{3}{2}\alpha J\frac{\gamma\cos\phi}{\sqrt{2J}}. \tag{10}$$

The system (10) is Hamiltonian with Hamiltonian energy

$$H(J,\phi)=-\Omega J+\frac{3}{4}\alpha J^2-\gamma\sqrt{2J}\cos\phi. \tag{11}$$

Suppose $\alpha>0$, $\Omega>0$ and $0<\gamma<\frac{4}{9}\sqrt{\Omega^3/\alpha}$. Then there exist centers $(J,\phi)=(j_1,\pi)$, $(j_3,0)$ and a hyperbolic saddle (j_2,π), where $0<j_1<j_2<j_3$ are roots of the cubic equation

$$x^3-2\left(\frac{2\Omega}{3\alpha}\right)x^2+\left(\frac{2\Omega}{3\alpha}\right)^2x-2\left(\frac{\gamma}{3\alpha}\right)^2=0. \tag{12}$$

Let $\mathbf{p}_i$, $i=1,2,3$, be the fixed points (j_1,π), (j_2,π), $(j_3,0)$. The level set

$$H(J,\phi)=H(J_2,\pi)\equiv H_0, \tag{13}$$

is composed of two homoclinic orbits Γ_+, Γ_- and the saddle $\mathbf{p}_2$. The homoclinic orbits $(J_\pm(t),\ \phi_\pm(t))$ are given by

$$J_\pm(t)=\pm\frac{2r_+r_-}{(r_+-r_-)\cosh at\pm(r_++r_-)}+j_2,$$
$$\phi_\pm(t)=\arcsin\left(\frac{-\dot{J}_\pm(t)}{\gamma\sqrt{2J_\pm(t)}}\right), \tag{14}$$

where $a=3\alpha\sqrt{-r_+r_-}$, $r_\pm=2(\kappa\pm\sqrt{2\kappa j_2})$, and $\kappa=2\Omega/3\alpha-j_2$. The phase portraits are shown in Fig. 1. Note that the phase space $(J,\ \phi)$ is considered as $\mathbf{R}^1\times S^1$. Also, the level set $H=0$ contains the line $J=0$ on which the vector field is singular, and thus the orbits with $H=0$ have discontinuities at $(J,\phi)=\left(0,\frac{1}{2}\pi\right)$ and $\left(0,\frac{3}{2}\pi\right)$.

For small $\mu>0$ the averaged system (9) may possess hyperbolic periodic orbits near fixed points $\mathbf{p}_i$, $i=1,2,3$, of equation (10). By the transformation (5) and the averaging theorem, these hyperbolic periodic orbits of equation (9) correspond to normally hyperbolic invariant 1-tori T_i, $i=1,2,3$, of the Poincaré map $P_{\epsilon,\mu}$ of equation (2) (cf., Yagasaki et al., 1990). Since the action of $P_{\epsilon,\mu}$ gives rise to the rotation by $\epsilon\nu T$ in the τ-direction, the invariant tori T_i, $i=1,2,3$, are not subjected to phase locking. Hence, the orbits on these invariant tori indicate quasi-periodic motions. The existence and stability of such

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

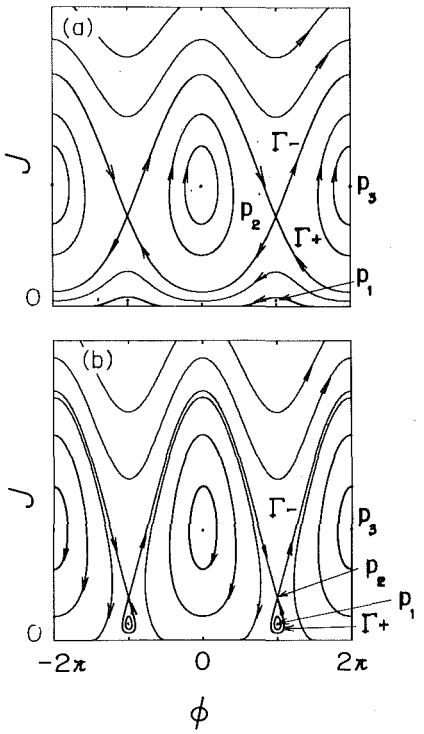


Fig. 1 Phase portraits of equation (10); (*a*) $0 < \gamma < 2\sqrt{2\Omega^3/\alpha}/9$, (*b*) $2\sqrt{2\Omega^3/\alpha}/9 < \gamma < 4\sqrt{\Omega^3/\alpha}/9$

quasi-periodic motions can be studied by the approach of Yagasaki et al. (1990). However, the same analysis is not repeated here. In the following we concentrate on the chaotic dynamics of the averaged system (9) and the weakly nonlinear oscillator (2).

3 Chaos in the Averaged System: Melnikov's Analysis

For $\mu=0$ the averaged system (9) has a saddle point $\mathbf{p}_2$ whose stable and unstable manifolds coincide, as shown in Fig. 1. For $\mu>0$ this coincidence is broken up and the stable and unstable manifolds may intersect transversely. If there exist such intersections, then chaotic motions of the Smale horseshoe type may occur. Melnikov's method provides an analytical tool to test for the existence of such chaotic motions. In this section Melnikov's method is applied to the averaged system (9). The reader should consult Guckenheimer and Holmes (1983) and Wiggins (1988b) for details on Melnikov's method.

As is standard, the Poincaré map $\bar{P}_\mu$ associated with the averaged system (9) is defined as

$$\bar{P}_\mu:(J(0),\phi(0))\to(J(\bar{T}),\phi(\bar{T})), \tag{15}$$

where $(J(t),\phi(t))$ is a solution of equation (9) and $\bar{T}=2\pi/\nu_0$. The unperturbed Poincaré map $\bar{P}_0$ has a hyperbolic fixed point $\mathbf{p}_2$ with coincident stable and unstable manifolds $W^s(\mathbf{p}_2)$, $W^u(\mathbf{p}_2)$. For $\mu>0$ sufficiently small $\bar{P}_\mu$ still possesses a hyperbolic fixed point $\mathbf{p}_2^\mu=\mathbf{p}_2+0(\mu)$ and its stable and unstable manifolds $W^s(\mathbf{p}_2^\mu)$, $W^u(\mathbf{p}_2^\mu)$ are close to $W^s(\mathbf{p}_2)$ and $W^u(\mathbf{p}_2)$, respectively, while they may not coincide.

The main idea of Melnikov's method is to follow $W^S(\mathbf{p}_2^\mu)$ and $W^u(\mathbf{p}_2^\mu)$ and measure the separation between them. The calculations are performed by using first-order variational equations around the unperturbed homoclinic orbits $(J_\pm(t), \phi_\pm(t))$. The distance $d_\pm(t_0)$ between $W^S(\mathbf{p}_2^\mu)$ and $W^u(\mathbf{p}_2^\mu)$ near $(J_\pm(t_0), \phi_\pm(t_0))$ (cf., Fig. 2) can be written as

$$d_\pm(t_0)=\mu K_\pm(t_0)^{-1}M_\pm(t_0)+0(\mu^2), \tag{16}$$

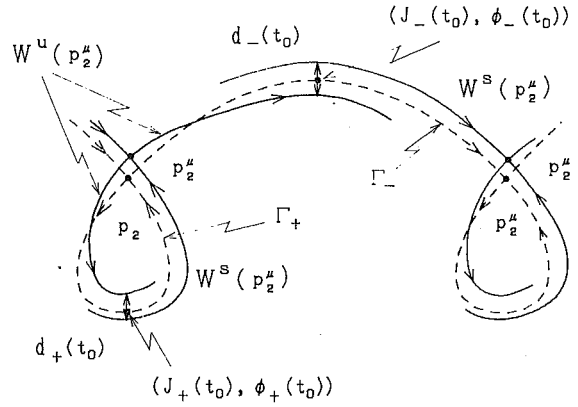


Fig. 2 Perturbed manifolds and the distance functions for case (*b*) in Fig. 1

where the subscripts + and −, respectively, indicate the lower and upper branches of these invariant manifolds; $M_\pm(t_0)$ are the Melnikov functions and $K_\pm(t_0)$ are nonzero 0(1) quantities. For the present example, $M_\pm(t_0)$ and $K_\pm(t_0)$ are given by

$$\begin{aligned}
M_\pm(t_0)=\int_{-\infty}^{\infty}\Bigg\{&\left[-\gamma\sqrt{2J_\pm(t)}\sin\phi_\pm(t)\right]\\
&\cdot\frac{\beta}{2}\cos[\nu_0(t-t_0)-2\phi_\pm(t)]\\
&-\left[-\Omega+\frac{3}{2}\alpha J_\pm(t)-\frac{\gamma\cos\phi_\pm(t)}{\sqrt{2J_{\pm(t)}}}\right]\\
&\times(-2\delta\omega J_\pm(t)-\beta J_\pm(t)\sin[\nu_0(t-t_0)-2\phi_\pm(t)])\Bigg\}dt,\\
=\beta\sin\nu_0 t_0\int_{-\infty}^{\infty}\Bigg\{&-\frac{1}{2}\gamma\sqrt{2J_\pm(t)}\sin\phi_\pm(t)\sin[\nu_0 t-2\phi_\pm(t)]\\
&-\left[-\Omega+\frac{3}{2}\alpha J_\pm(t)-\frac{\gamma\cos\phi_\pm(t)}{\sqrt{2J_\pm(t)}}\right]\cos[\nu_0 t-2\phi_\pm(t)]\Bigg\}dt\\
&+\delta\omega\cdot2\int_{-\infty}^{\infty}\left[-\Omega+\frac{3}{2}\alpha J_\pm(t)-\frac{\gamma\cos\phi_\pm(t)}{\sqrt{2J_\pm(t)}}\right]J_\pm(t)dt\\
=\beta\sin\nu_0 t_0&\cdot A_\pm(\alpha,\gamma,\Omega,\nu_0)+\delta\omega B_\pm(\alpha,\gamma,\Omega),
\end{aligned}\tag{17}$$

$$\begin{aligned}
K_\pm(t_0)=\Bigg\{&\left[-\gamma\sqrt{2J_\pm(t_0)}\sin\phi_\pm(t_0)\right]^2\\
&+\left[-\Omega+\frac{3}{2}\alpha J_\pm(t_0)-\frac{\gamma\cos\phi_\pm(t_0)}{\sqrt{2J_\pm(t_0)}}\right]^2\Bigg\}^{1/2}.
\end{aligned}\tag{18}$$

The Melnikov function $MI(t_0)$ has a simple zero if

$$\frac{\beta}{\delta\omega}>|B_\pm(\alpha,\gamma,\Omega)/A_\pm(\alpha,\gamma,\Omega,\nu_0)|=R_\pm(\alpha,\gamma,\Omega,\nu_0). \tag{19}$$

Since $K_\pm(t_0)\neq0$, we see from equation (16) that, if $M_\pm(t_0)$ has a simple zero at s, then by the implicit function theorem $d_\pm(t_0)$ also has a simple zero near s, and hence there exist transverse intersections between $W^s(\mathbf{p}_2^\mu)$ and $W^u(\mathbf{p}_2^\mu)$ which yield transverse homoclinic orbits. By the Smale-Birkhoff homoclinic theorem, the presence of such orbits implies that there exists an invariant Cantor set Λ on which $\bar{P}_\mu$ is equivalent to the Smale horseshoe map; orbits of $\bar{P}_\mu$ on Λ correspond to realizations of a completely random process, coin tossing, and

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

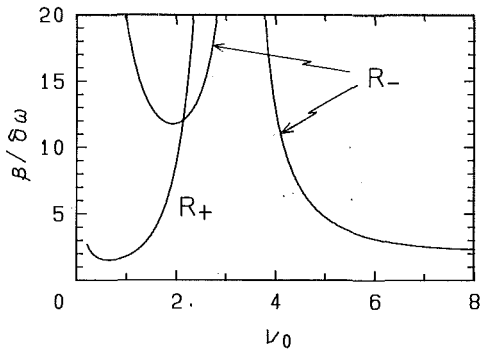


Fig. 3 Homoclinic bifurcation curves, $\alpha = 2$, $\gamma = 1.5$, $\Omega = 3$

Λ contains a countable set of periodic orbits of arbitrarily long periods, an uncountable set of bounded nonperiodic orbits, and a dense orbit (cf., Guckenheimer and Holmes, 1983). Thus, condition (19) gives the regions on parameter space where chaotic dynamics may occur in the averaged system (9).

Here we remark that the invariant set Λ is not an attractor, and hence condition (19) does not guarantee the existence of a strange attractor, i.e., steady-state chaos. However, if there exists such an invariant set, one can observe at least transient chaotic motions; they eventually settle into stable periodic motions, but look chaotic for long time. Moreover, if there are two or more attractors, the existence of transverse homoclinic orbits implies that fractal basin boundaries exist and unpredictability in the presence of uncertainties in parameters occurs (cf., Moon, 1987). Thus, the dynamics of P_μ is very complicated when condition (19) is met.

A numerical example was computed. The parameters were chosen as follows: $\alpha = 2$, $\gamma = 1.5$, and $\Omega = 3$ with $\delta\omega$, β, and ν_0 varying. We note that $2\sqrt{2\Omega^3/\alpha}/9 < \gamma < 4\sqrt{\Omega^3/\alpha}/9$ and the unperturbed system (10) has a phase portrait such as shown in Fig. 1(*b*). The integral $B_\pm(\alpha,\gamma,\Omega)$ can be easily evaluated to yield

$$B_+(\alpha,\gamma,\Omega) = -16\Omega\theta_0/3\alpha + 6\rho, \tag{20a}$$

$$B_-(\alpha,\gamma,\Omega) = 16\Omega(\pi-\theta_0)/3\alpha + 6\rho, \tag{20b}$$

where $\theta_0 = \arccos\left(\dfrac{r_+ + r_-}{r_+ - r_-}\right)$ and $\rho = \sqrt{-r_+ r_-}$, which depend only on α, γ, and Ω. In this example $\theta_0 = \dfrac{\pi}{4}$ and $\rho = 1$, and hence $B_+ = -2\pi + 6$ and $B_- = 6\pi + 6$. Since the integral $A_\pm(\alpha,\gamma,\Omega,\nu_0)$ is difficult to evaluate, the "tails" of the improper integrals were cut off and the normal integrals were numerically computed by using Simpson's rule. Note that the integrand tends to zero exponentially as $t \to \pm\infty$. In this calculation there are two kinds of errors associated with the missing "tails" and the integrator. These errors can be estimated and their bounds obtained.

Figure 3 shows the homoclinic bifurcation curves $\beta/\delta\omega = R_\pm(\alpha,\gamma,\Omega,\nu_0)$ in $\beta/\delta\omega - \nu_0$ space. In this case three types of chaotic motions are expected. In the region $\beta/\delta\omega > R_+$, there exist motions which erratically oscillate around $\phi = \pi$, roughly along Γ_+. In the region $\beta/\delta\omega > R_-$, there exist motions which erratically rotate in the ϕ-direction, roughly along Γ_-. In the region where $\beta/\delta\omega > R_+$ and $\beta/\delta\omega > R_-$, besides these two types of chaotic motions, there exist motions which oscillate around $\phi = \pi$ and rotate approximately 2π in the ϕ-direction in arbitrary orders.

4 Chaos in the Weakly Nonlinear Oscillator

In the preceding section, using the Melnikov method, we calculated conditions on parameters for which the averaged system (9) has transverse homoclinic orbits yielding chaotic dynamics. The existence of such orbits in the averaged system (9) also implies that the full system (6) or the original system (2) has a normally hyperbolic invariant set with a Cantor-set-like structure. This is proved through the lines given by Holmes (1986) and Yagasaki et al. (1990). In this section, we describe the dynamics of the weakly nonlinear oscillator (2) on such an invariant set by a different approach.

Consider systems of the general form

$$\dot{\mathbf{x}} = \epsilon\mathbf{f}(\mathbf{x},\epsilon\nu t,\omega t), \tag{21}$$

where

$$\mathbf{x} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in \mathbf{R}^2, \tag{22}$$

and $\mathbf{f}(\mathbf{x},\tau,\theta)$ is 2π-periodic in both τ and θ. Note that equation (6) is of the form (21). By introducing the new state variable $\tau = \epsilon\nu t$, equation (21) can be rewritten as

$$\dot{\mathbf{x}} = \epsilon\mathbf{f}(\mathbf{x},\tau,\omega t), \quad \dot{\tau} = \epsilon\nu. \tag{23}$$

Applying the averaging method to equation (23), one obtains

$$\dot{\mathbf{y}} = \epsilon\bar{\mathbf{f}}(\mathbf{y},\tau), \quad \dot{\tau} = \epsilon\nu, \tag{24}$$

where

$$\bar{\mathbf{f}}(\mathbf{y},\tau) = \frac{1}{T}\int_0^T \mathbf{f}(\mathbf{y},\tau,\omega t)\,dt, \tag{25}$$

with $T = 2\pi/\omega$. The averaging theorem then implies that if $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are, respectively, solutions of (23) and (24) with $\mathbf{x}(0) = \mathbf{y}(0) + 0(\epsilon)$, then

$$\mathbf{x}(t) = \mathbf{y}(t) + 0(\epsilon), \tag{26}$$

on a time scale $t \sim 1/\epsilon$.

We suspend (23) and (24) over the phase space to $\mathbf{R}^2 \times S^1 \times S^1$ as follows:

$$\dot{\mathbf{x}} = \epsilon\mathbf{f}(\mathbf{x},\tau,\theta), \quad \dot{\tau} = \epsilon\nu, \quad \theta = \omega, \tag{27}$$

$$\dot{\mathbf{y}} = \epsilon\bar{\mathbf{f}}(\mathbf{y},\tau), \quad \dot{\tau} = \epsilon\nu, \quad \theta = \omega. \tag{28}$$

Define a global cross-section $\Sigma_\tau = \{(\mathbf{x},\tau,\theta) \in \mathbf{R}^2 \times S^1 \times S^1 \mid \tau = 0\}$. Then, Poincaré maps Φ_ϵ, $\Phi_0 : \Sigma_\tau \to \Sigma_\tau$ generated by the flow of (27) and (28) are given by

$$\Phi_\epsilon : (\mathbf{x}(0),\theta) \to (\mathbf{x}(T_1/\epsilon),\theta + \omega T_1/\epsilon), \tag{29}$$

$$\Phi_0 : (\mathbf{y}(0),\theta) \to (\mathbf{y}(T_1/\epsilon),\theta + \omega T_1/\epsilon), \tag{30}$$

where $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are solutions of (23) and (24), and $T_1 = 2\pi/\nu$. It follows from the estimation (26) that Φ_ϵ is well approximated by Φ_0 as

$$\Phi_\epsilon = \Phi_0 + 0(\epsilon). \tag{31}$$

By letting $\bar{\Sigma} = \{(\mathbf{y},\tau) \in \mathbf{R}^2 \times S^1 \mid \tau = 0\}$ be a global cross-section, the Poincaré map $F_0 : \bar{\Sigma} \to \bar{\Sigma}$ for the flow of (24) is also obtained by

$$F_0 : \mathbf{y}(0) \to \mathbf{y}(T_1/\epsilon). \tag{32}$$

Then, one has the relationship between Φ_0 and F_0 as follows:

$$\Phi_0(\mathbf{y},\theta) = (F_0(\mathbf{y}),\theta + \omega T_1/\epsilon). \tag{33}$$

Suppose F_0 has a hyperbolic fixed point $\mathbf{p}_0$ whose stable and unstable manifolds $W^S(\mathbf{p}_0)$, $W^u(\mathbf{p}_0)$ intersect transversely. This means that there exist transverse homoclinic orbits in the averaged system (24). It follows from equation (33) that Φ_0 has a normally hyperbolic invariant 1-torus $\gamma_0 = \mathbf{p}_0 \times S^1$ and its stable and unstable manifolds $W^S(\gamma_0)$, $W^u(\gamma_0)$ intersect transversely in a torus, as shown in Fig. 4. Then, it is shown by the invariant manifold theorem (Hirsch et al., 1977) that $\Phi_\epsilon = \Phi_0 + 0(\epsilon)$ has a normally hyperbolic invariant 1-torus γ_ϵ near γ_0 and its stable and unstable manifolds $W^s(\gamma_\epsilon)$, $W^u(\gamma_\epsilon)$ intersect transversely in a torus like $W^s(\gamma_0)$ and $W^u(\gamma_0)$. Thus, there exist orbits homoclinic to the normally hyperbolic invariant 1-torus γ_ϵ.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

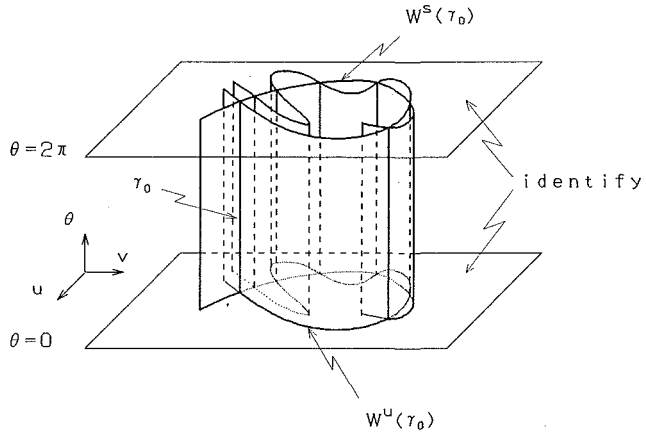


Fig. 4 Stable and unstable manifolds for the averaged Poincaré map Φ_0

In this situation the Smale-Birkhoff homoclinic theorem does not apply since it describes the mechanism only for chaos arising from the presence of orbits homoclinic to hyperbolic periodic orbits. Wiggins (1988b) generalized this theorem to the case of orbits homoclinic to normally hyperbolic invariant tori. By applying his results, it is shown that the dynamics of Φ_ϵ normal to the invariant torus γ_ϵ is chaotic. We now describe this type of chaotic dynamics.

Let $B \subset \mathbf{R}^2 \times S^1$ be a small region near γ_ϵ, partially bounded by a piece of $W^u(\gamma_\epsilon)$, as shown in Fig. 5(*a*). Then, the existence of orbits homoclinic to γ_ϵ implies that for some iterate of the Poincaré map Φ_ϵ^N, the image of B is a horseshoe shape and intersects B. Thus, two slabs (rings) H_1, H_2 can be found whose images V_1, V_2 intersect H_1, H_2 (cf., Fig. 5(*b*)). Here we define a set Ξ as

$$\Xi = \bigcap_{n=-\infty}^{\infty} (\Phi_\epsilon^N)^n(B). \tag{34}$$

The set Ξ is nonempty and consists of points which never leave B as Φ_ϵ^N is iterated. Thus, Ξ is an invariant set of Φ_ϵ^N.

We associate a point $(\mathbf{x},\theta) \in \Xi$ to a bi-infinite sequence $\mathbf{a} = \{a_i \mid -\infty < i < \infty\}$:

$$a_i = \begin{cases} 1 & \text{if } (\Phi_\epsilon^N)^i(\mathbf{x},\theta) \in H_1, \\ 2 & \text{if } (\Phi_\epsilon^N)^i(\mathbf{x},\theta) \in H_2. \end{cases} \tag{35}$$

Let E be the set of bi-infinite sequences of two symbols $\{1,2\}$ and σ the shift map $\sigma: E \to E$, which takes a sequence and shifts indices one place, i.e., $\{\sigma(\mathbf{a})\}_i = a_{i+1}$. For the correspondence h between Ξ and E defined by equation (35), one has

$$h \circ \Phi_\epsilon^N|_\Xi = \sigma \circ h, \tag{36}$$

like the Smale horseshoe map. Note that σ has a countably infinite set of periodic orbits such as $\{\ldots, 1, 2, 1, 2, \ldots\}$, $\{\ldots, 1, 2, 2, 1, 2, 2, \ldots\}$, ..., an uncountably infinite set of nonperiodic orbits, and a dense orbit in metric defined by

$$d(\mathbf{a},\mathbf{b}) = \sum_{i=-\infty}^{\infty} \delta_i \, 2^{-|i|}, \quad \delta_i = \begin{cases} 0 & \text{if } a_i = b_i, \\ 1 & \text{if } a_i \neq b_i, \end{cases} \tag{37}$$

where $\mathbf{a} = \{a_i\}$ and $\mathbf{b} = \{b_i\}$. Since h is not a 1-1 map, an orbit of σ does not correspond to an orbit of Φ_ϵ^N. However, one can associate a sequence $\mathbf{a} \in E$ to a torus $\{(\mathbf{x}^\theta,\theta) \mid \theta \in S^1\}$ in Ξ such that $h(\Phi_\epsilon^N)^i \mathbf{x}^\theta = a_i$, $i = 1,2,\ldots$, for all $\theta \in S^1$. Periodic, nonperiodic, and a dense orbit in E, respectively, correspond to finite, infinite, and a dense set of tori in Ξ. The invariant set Ξ consists of infinitely many tori and has Cantor-type structure. Thus, the dynamics of Φ_ϵ^N restricted to Ξ in directions normal to these tori, i.e., in the $\mathbf{x}$-plane, is "Smale-horseshoe-like," and hence

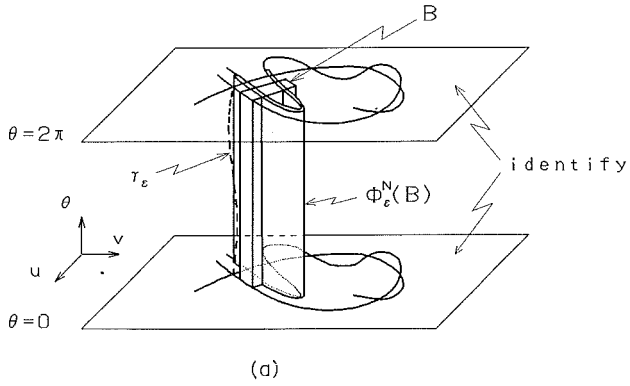


Fig. 5(*a*) Construction of horseshoe (u, v, θ) space

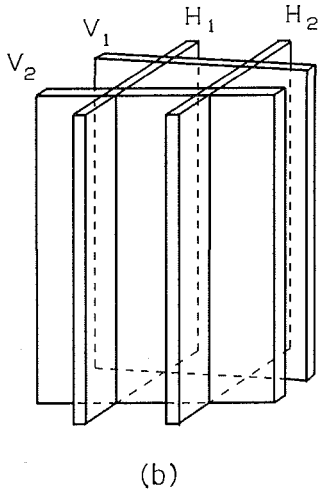


Fig. 5(*b*) Two slabs H_1, H_2

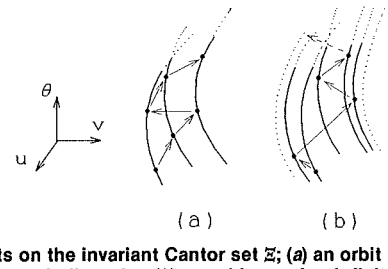


Fig. 6 Orbits on the invariant Cantor set Ξ; (*a*) an orbit passing finitely many tori in a periodic order, (*b*) an orbit passing infinitely many tori in an erratic order

chaotic. In Ξ there exist (i) orbits passing finitely many tori in periodic orders as shown in Fig. 6(*a*), (ii) orbits passing infinitely many tori in arbitrary orders as shown in Fig. 6(*b*), and (iii) orbits passing infinitely many tori which construct a dense set in Ξ. These orbits only rotate by $N\omega T/\epsilon$ in the θ-direction.

We next turn to the dynamics of the original system (2). Suppose condition (19) is met. Then it follows from the Melnikov analysis in Section 3 that there exist orbits homoclinic to a hyperbolic periodic orbit in the averaged system (7). As described above, this implies that chaos may occur in the full system (6). Since $J(t)$ and $\phi(t)$, respectively, represent the slowly varying amplitude and phase by equation (5), one can expect that the response of the weakly nonlinear oscillator (2) indicates chaotic motions with randomlike modulations of amplitude and phase. Thus, equation (19) also gives the regions in parameter space where chaotic dynamics may exist in the weakly nonlinear oscillator (2).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

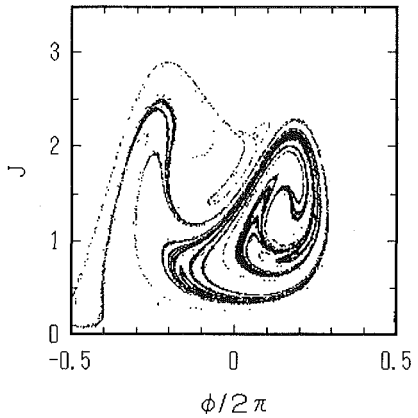


Fig. 7 A numerically computed orbit of the Poincaré map $\bar{P}_\mu$ of the averaged system (7), $\alpha = 2$, $\gamma = 1.5$, $\Omega = 3$, $\nu_0 = 1.8$, $\mu\beta = 1.5$, $\mu\delta\omega = 0.1$

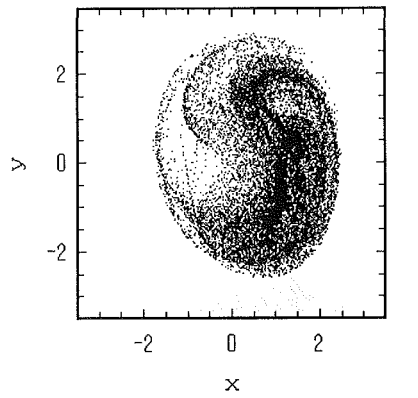


Fig. 8 A numerically computed orbit of the Poincaré map $P_{\epsilon,\mu}$ of (2) onto (x,y)-plane, $\bar{\alpha} = 0.2$, $\bar{\gamma} = 0.15$, $\omega = 1.14$, $\bar{\nu} = 2.359$, $\bar{\beta} = 0.15$, $\bar{\delta} = 0.008772$

5 Numerical Simulation

Numerical integrations of the averaged system (9) and the original system (2) were performed by the Runge-Kutta-Gill method and sustained chaotic motions were observed for some parameter values.

Figure 7 shows a strange attractor of the Poincaré map $\bar{P}_\mu$ of the averaged system (9) for $\alpha = 2$, $\gamma = 1.5$, $\Omega = 3$, $\nu_0 = 1.8$, $\mu\beta = 1.5$, $\mu\delta\omega = 0.1$. This orbit was computed by transforming (9) into a differential equation and solving it numerically since we could not sometimes integrate (9) directly because of the singular property about $J = 0$ (see Yagasaki et al.,1990).

Figure 8 shows a projection of the computed orbit for the Poincaré map $P_{\epsilon,\mu}$ of equation (2) onto the (x, y)-plane for $\bar{\alpha} = \epsilon\alpha = 0.2$, $\bar{\gamma} = \epsilon\gamma = 0.15$, $\omega = 1.14$, $\bar{\nu} = 2\omega + \epsilon\nu = 2.359$, $\bar{\beta} = \epsilon\mu\beta = 0.15$, $\bar{\delta} = \epsilon\mu\delta = 0.008772$. These parameter values correspond to those of Fig. 7 when $\epsilon = 0.1$. Figure 9 shows a double Poincaré section (cf., Moon, 1987) for the strange attractor displayed in Fig. 8. This figure was drawn by taking the thin section $0 < \tau < 0.02 \times 2\pi$ and projecting points falling in this slice onto the (x, y)-plane. A fractallike structure of the strange attractor in Fig. 8 is observed in Fig. 9. Figure 10 shows a numerical solution $x(t)$ of equation (2) for these parameter values.

6 Conclusions

In this paper we have studied the chaotic dynamics of a weakly nonlinear oscillator with parametric and external resonances. By using the averaging method and Melnikov's technique, the existence of orbits homoclinic to a normally hyperbolic invariant torus was detected and a criterion for the existence of chaos was obtained. The mechanism and structure of chaos in quasi-periodically forced weakly nonlinear oscillators were described. Numerical simulation results also demonstrated the existence of sustained chaotic motions.

Weakly nonlinear two degree-of-freedom systems with periodic forcing are frequently transformed into four-dimensional autonomous systems with one angular variable by using the averaging technique and neglecting higher-order terms. Chaos with the structure described here, may occur in such two degree-of-freedom systems. The existence of this type of chaotic dynamics can be also detected by a version of Melnikov's technique due to Holmes and Marsden (1982) or Wiggins and Holmes (1987). See also Shaw and Wiggins (1988) and Wiggins and Shaw (1988). Holmes (1986) performed such an analysis for a two-mode model of surface waves in containers although he did not describe the mechanism for chaos in the periodically forced two degree-of-freedom system.

The system considered here may be a little specialized, particularly in the scaling of the parameters. However, our results suggest that in a large class of weakly nonlinear single degree-of-freedom systems with two-frequency forcing near simultaneous resonance, chaotic dynamics may exist for wide ranges of the parameters. This is a contrast to the case of single frequency force in which for nonzero damping chaos may occur only in exponentially small wedges in parameter space (see Guckenheimer and Holmes, 1983; Holmes et al., 1987).

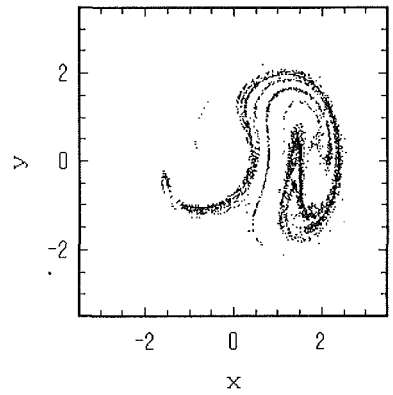


Fig. 9 Double Poincaré section, $\bar{\alpha} = 0.2$, $\bar{\gamma} = 0.15$, $\omega = 1.14$, $\bar{\nu} = 2.359$, $\bar{\beta} = 0.15$, $\bar{\delta} = 0.008772$

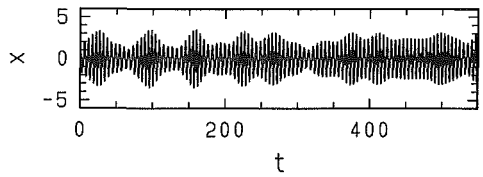


Fig. 10 A numerical solution of equation (2), $\bar{\alpha} = 0.2$, $\bar{\gamma} = 0.15$, $\omega = 1.14$, $\bar{\nu} = 2.359$, $\bar{\beta} = 0.15$, $\bar{\delta} = 0.008772$

Acknowledgments

The author would like to thank Profs. N. Yanagihara, Y. Hamamatsu, and H. Kamegaya of Tamagawa University for helpful suggestions and encouragement. He also acknowledges one of the reviewers for useful comments and constructive suggestions.

References

Guckenheimer, J., and Holmes, P. J., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Hale, J. K., 1969, *Ordinary Differential Equations*, John Wiley and Sons, New York.

HaQuang, N., Mook, D. T., and Plaut, R. H., 1987, "A Nonlinear Analysis of the Interactions Between Parametric and External Excitations," *Journal of Sound and Vibration*, Vol. 118, pp. 425–439.

Hirsch, M. W., Pugh, C. C., and Shub, M., 1977, *Invariant Manifolds*, (Lecture Notes in Mathematics, No. 583), Springer-Verlag, New York.

Holmes, P. J., 1986, "Chaotic Motions in a Weakly Nonlinear Model for Surface Waves," *Journal of Fluid Mechanics*, Vol. 162, pp. 365–388.

Holmes, P. J., and Marsden, J. E., 1982, "Horseshoes in Perturbations of Hamiltonian Systems With Two Degree of Freedom," *Communications in Mathematical Physics*, Vol. 82, pp. 523–544.

Holmes, P., Marsden, J., and Scheurle, J., 1987, "Exponentially Small Splitting of Separatrices," *Proceedings of the National Academy of Sciences of the United States of America*, to be published.

Ide, K., and Wiggins, S., 1989, "The Bifurcation to Homoclinic Tori in the Quasiperiodically Forced Duffing Oscillator," *Physica D*, Vol. 34, pp. 169–182.

Melnikov, V. K., 1963, "On the Stability of the Center for Time Periodic Perturbations," *Transactions of the Moscow Mathematical Society*, Vol. 12, pp. 1–56.

Moon, F. C., 1987, *Chaotic Vibrations, an Introduction for Applied Scientists and Engineers*, John Wiley and Sons, New York.

Ness, D. J., 1971, "Resonance Classification in a Cubic System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, pp. 585–590.

Sanders, J. A., and Verhulst, F., 1985, *Averaging Methods in Nonlinear Dynamical Systems*, Springer-Verlag, New York.

Scheurle, J., 1986, "Chaotic Solutions of Systems With Almost Periodic Forcing," *Journal of Applied Mathematics and Physics*, Vol. 37, pp. 12–26.

Shaw, S. W., and Wiggins, S., 1988, "Chaotic Dynamics of a Whirling Pendulum," *Physica D*, Vol. 31, pp. 190–211.

Troger, H., and Hsu, C. S., 1977, "Response of a Nonlinear System Under Combined Parametric and Forcing Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 44, pp. 179–181.

Wiggins, S., 1988a, "On the Detection and Dynamical Consequences of Orbits Homoclinic to Hyperbolic Periodic Orbits and Normally Hyperbolic Invariant Tori in a Class of Ordinary Differential Equations," *SIAM Journal of Applied Mathematics*, Vol. 48, pp. 262–285.

Wiggins, S., 1988b, *Global Bifurcations and Chaos—Analytical Methods*, Springer-Verlag, New York.

Wiggins, S., and Holmes, P., 1987, "Homoclinic Orbits in Slowly Varying Oscillators," *SIAM Journal on Mathematical Analysis*, Vol. 18, pp. 612–629; Errata, Vol. 19, pp. 1254–1255.

Wiggins, S., and Shaw, S. W., 1988, "Chaos and Three-Dimensional Horseshoes in Slowly Varying Oscillators," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 959–968.

Yagasaki, K., Sakata, M., and Kimura, K., 1990, "Dynamics of a Weakly Nonlinear System Subjected to Combined Parametric and External Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp. 209–217.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A Modified, Slowly Varying Parameter Approach for Systems With Impulsive Loadings

Nicholas P. Jones
Assistant Professor,
Department of Civil Engineering,
The Johns Hopkins University,
Baltimore, MD 21218-2699

Harry W. Shenton III[1]
Research Structural Engineer,
National Institute of Standards and Technology,
Gaithersburg, MD 20899

Many oscillating systems in practice are subject to periodic, impulsive loads during their motion. For nonlinear systems traditional methods, such as the slowly varying parameter technique, have been used to study the response of such systems by "smearing" the effect of the impulse over one cycle of the oscillation. Unfortunately, in many instances this approach violates the basic assumption used in the application of the averaging technique, resulting in limited predictive capability. This paper presents a modification to the standard slowly varying parameter technique for nonlinear systems which addresses this problem by assuming a solution form which admits a discontinuity in velocity at the time the impulse is applied. Examples are given which outline the improvement in accuracy of the new technique.

Introduction

Many physical systems are subjected to periodic loadings which are impulsive in nature, i.e., the force is applied over a time interval which is much shorter than a characteristic time associated with the resulting motion (usually the period). Common examples are the mechanical clock (Butenin, 1965), the R-C circuit (Samoilenko, 1961), the rocking block (Spanos and Koh, 1984) and the impact damper (Masri and Caughey, 1966). Traditionally, these loads are considered to produce a discontinuity in velocity of appropriate magnitude. In linear vibrating systems this discontinuity is relatively straightforward to handle: a piecewise-continuous solution in velocity may be obtained, which represents the sum of transient and periodic solutions between the impulses, with a velocity condition imposed when the impulse occurs. In nonlinear systems, however, this approach becomes quite difficult, and it is not possible, in general, to produce the piecewise-exact solution.

Several widely known methods are available for the approximate solution of mildly nonlinear systems to periodic loading. These include Galerkin's method (Nayfeh, 1979), equivalent linearization (Caughey, 1963; Iwan, 1969), harmonic balance (Nayfeh, 1979), and the method of slowly varying parameters[2] (Krylov and Bogoliubov, 1970). The fundamental steady-state solution is identical in each of these methods. All essentially assume that the nonlinearity is "small enough" so that an approximate harmonic solution may be obtained by averaging the nonlinear effect over one period of the resulting motion.

A number of investigators have applied such methods for prediction of the response of nonlinear systems subjected to impulsive loadings. Samoilenko (1961) and Butenin (1965) used a slowly varying parameter approach in which the periodic impulse was represented as a delta function in force; Spanos and Koh (1984) used equivalent linearization to extract an effective linear stiffness and damping in order to estimate the response of a rocking rigid block to harmonic ground motion.

The problem in the application of the aforementioned methods to impulsively loaded systems is that the basic averaging assumption used is violated by the presence of the impulse. For example, in the slowly varying parameter approach, it can no longer be assumed that the amplitude and phase are slowly varying, due to the discontinuity in velocity resulting from the impulse. Smearing this effect over one cycle of oscillation is a rather coarse approximation of the actual behavior. In cases where the maximum amplitude is of interest, or a more accurate representation of the motion throughout the period is required, application of the standard averaging techniques does not appear appropriate.

More recently, Samoilenko (1967, 1974) and Mitropolsky and Samoilenko (1985) proposed a modified solution form using the sum of a continuous, slowly varying function and an infinite series which captured the discontinuity in the response. The solution was developed specifically for systems with one impulse per cycle.

This paper presents an alternative method for solving systems subjected to such impulsive loads. The technique is based on the method of slowly varying parameters, with a modified assumed form of solution. The effect of the impulsive action is included in a relatively simple form, with the resulting solution in the first approximation characterizing more realist-

[1]Formerly: Graduate Student, Department of Civil Engineering, The Johns Hopkins University, Baltimore, MD 21218-2699.

[2]Also known as the method of averaging or the Krylov-Bogoliubov (KB) Method.

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 23, 1989; final revision, Nov. 3, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

ically the actual response during steady-state motion. At present, the development is restricted to symmetric, periodic solutions at the frequency of excitation.

Stability of the solution form will not be considered in this paper. A discussion of this topic, which becomes quite involved, will be given in a subsequent paper.

The modified steady-state solution is first presented for direct comparison with the standard harmonic solution. The method of slowly varying parameters is then applied using this form of solution, and a modified set of equations is derived from which the approximate amplitude and phase of the motion can be computed. The technique is demonstrated using two examples: a single degree-of-freedom system with impulses and a rocking block subject to harmonic ground motion. Discussion of the solution form in the context of these results is given.

While results are presented here for two examples where piecewise-exact solutions are available, the technique has much broader application. For many nonlinear single and multiple degree-of-freedom systems with impulses, the piecewise-exact solutions are not available. Solutions are attainable only through an averaging technique, such as the method of slowly varying parameters. The method proposed herein offers an alternative solution for such problems which is efficient and accurate.

The Modified Steady-State Solution

Consider the response of a system governed by the equation of motion

$$\ddot{x} + \omega_n^2 x = \mu f(x,\dot{x}) + B\cos\omega t \tag{1}$$

and the impulse-momentum condition

$$\dot{x}^- + I(x, \dot{x}^-) = \dot{x}^+. \tag{2}$$

In the foregoing equations, μ is a small dimensionless parameter, $I(x, \dot{x}^-)$ an impulse per unit mass applied at x, and $\dot{x}^-$ and $\dot{x}^+$ the velocity immediately before and immediately after the impulse is applied, respectively. In the following development, impulses are restricted to occur at $x=0$.

In the special case that $f(x, \dot{x})$ admits an exact solution to equation (1) between impulses (e.g., $f(x, \dot{x})$ is linear), a piecewise exact solution to equations (1) and (2) can be obtained; otherwise, an approximate solution is sought. To this end, investigators have relied on the classical method of slowly varying parameters (SVP) to find solutions to such systems. However, a modification to the standard technique was necessary to include the effect of the impulse-momentum condition: impulsive forces were applied to the right-hand side of the equation of motion with a delta function (e.g., Samoilenko, 1961; Butenin 1965),

$$\ddot{x} + \omega_n^2 x = \mu f(x,\dot{x}) + I(x,\dot{x}^-)|\dot{x}|\delta(x) + B\cos\omega t. \tag{3}$$

At this point, the SVP solution technique proceeded as usual using the assumed form of solution:

$$x(t) = A(t)\cos[\omega t - \psi(t)] \tag{4}$$

in which $A(t)$ and $\psi(t)$ represented a slowly varying amplitude and phase, respectively. It should be noted that under steady-state conditions, equation (4) reduces to

$$x(t) = A_0\cos[\omega t - \psi_0] \tag{5}$$

in which A_0 and ψ_0 denote the steady-state amplitude and phase, respectively. It is clear that equation (5) and its derivative cannot accurately represent the discontinuity occurring at $x=0$ as a result of the impulse.

For the system described by equations (1) and (2), a symmetric, periodic solution may exist at the frequency of excitation if two impulses, equal in magnitude but alternating in sign, occur each cycle. Note that under steady-state conditions, the magnitude of impulse applied is constant, regardless of the form of I, i.e., $|I(x,\dot{x}^-)| = |I(\dot{x}^-)| = \text{constant}$. With this in mind, the following modified, *steady-state* form of solution is proposed

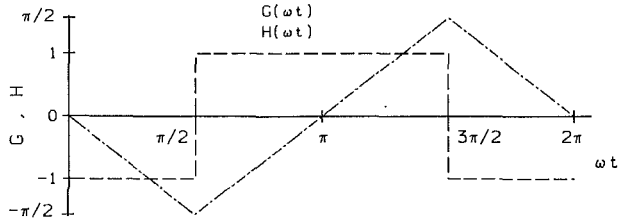


Fig. 1 Displacement $G(t)$ and velocity $H(t)$ modifying functions

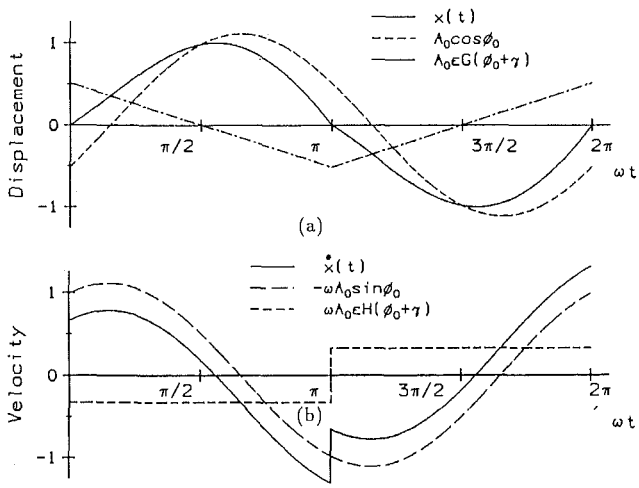


Fig. 2 Modified (*a*) displacement and (*b*) velocity profiles

$$x(t) = A_0\cos(\omega t - \psi_0) + A_0\epsilon G(\omega t - \psi_0 + \gamma). \tag{6}$$

Differentiating $x(t)$ with respect to t yields the *steady-state* velocity

$$\dot{x}(t) = -\omega A_0\sin(\omega t - \psi_0) + \omega A_0\epsilon H(\omega t - \psi_0 + \gamma). \tag{7}$$

In equations (6) and (7), A_0 and ψ_0 are referred to as the steady-state amplitude and phase; however, they do not represent the peak amplitude of response or total phase. $G(\omega t - \psi_0 + \gamma)$ is a periodic triangular wave with frequency ω and peak amplitude $\pi/2$, defined by

$$G(\omega t) = \begin{cases} -\omega t & ; \ 0 \le \omega t \le \pi/2 \\ \omega t - \pi & ; \ \pi/2 \le \omega t \le 3\pi/2 \\ -\omega t + 2\pi & ; \ 3\pi/2 \le \omega t \le 2\pi \end{cases} \tag{8}$$

as shown in Fig. 1; $H(\omega t - \psi_0 + \gamma)$ is a periodic square wave with frequency ω and unit amplitude, defined by

$$H(\omega t) = \begin{cases} -1 ; & 0 \le \omega t \le \pi/2 \\ +1 ; & \pi/2 \le \omega t \le 3\pi/2 \\ -1 ; & 3\pi/2 \le \omega t \le 2\pi \end{cases} \tag{9}$$

also shown in Fig. 1. Note that $dG/dt = \dot{G}(\omega t - \psi_0 + \gamma) = \omega H(\omega t - \psi_0 + \gamma)$.

The parameters ϵ and γ are determined such that during steady state, equation (2) is satisfied for the two impulses per cycle occurring at $x=0$. Note that two constraints exist in equation (2): one in velocity and one in displacement. Substituting equations (6) and (7) into equation (2) yields

$$\left.\begin{aligned} A_0\cos\phi_0|_{\phi_1} + A_0\epsilon G(\phi_0 + \gamma)|_{\phi_1} = 0 \\ A_0\cos\phi_0|_{\phi_2} + A_0\epsilon G(\phi_0 + \gamma)|_{\phi_2} = 0 \end{aligned}\right\} \tag{10}$$

$$\left.\begin{aligned} [-\omega A_0\sin\phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]_{\phi_1^-} + I(A_0, \phi_0, \epsilon, \gamma)_{\phi_1^-} \\ = [-\omega A_0\sin\phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]_{\phi_1^+} \\ [-\omega A_0\sin\phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]_{\phi_2^-} + I(A_0, \phi_0, \epsilon, \gamma)_{\phi_2^-} \\ = [-\omega A_0\sin\phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]_{\phi_2^+} \end{aligned}\right\} \tag{11}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

in which $\phi_0 = \omega t - \psi_0$, $\phi_1 = \omega t_1 - \psi_0$, $\phi_2 = \omega t_2 - \psi_0$ and t_1 and $t_2 = t_1 + \pi/\omega$ denote the two times during a period at which the impulses occur. As the solution is assumed symmetric, it is necessary to solve for ϵ and γ at only one time, for example $\phi = \phi_1$. Equations (10) and (11) then constitute two equations for the two unknowns, ϵ and γ.

Sample curves are presented in Fig. 2 for $x(t)$ and $\dot{x}(t)$ as given by equations (6) and (7) for representative values of $\epsilon = 0.295$, $\gamma = 27.64$ deg, $A_0 = 1.1093$, $\omega = 1.0$, and $\psi_0 = 117.64$ deg. The contribution from each component is indicated along with the total solution. Note that the modified solution contains an instantaneous change in slope of displacement and discontinuity in velocity, as desired.

Determination of the Steady-State Amplitude and Phase

To evaluate the steady-state amplitude (A_0) and phase (ψ_0), standard SVP procedures are followed using the modified solution form in equation (6). The amplitude and phase associated with the trigonometric term in equation (6) are assumed to be slowly varying functions of time, i.e.,

$$x(t) = A(t)\cos[\omega t - \psi(t)] + A_0\epsilon G(\omega t - \psi_0 + \gamma). \quad (12)$$

Note that the amplitude and phase associated with $G(t)$ take the steady-state values, *a priori*. Denoting $A(t) = A$, $\psi(t) = \psi$, $\phi = \omega t - \psi(t)$ and $\phi^* = -\psi + \psi_0 - \gamma$, equation (12) can be written as

$$x(t) = A\cos\phi + A_0\epsilon G(\phi - \phi^*). \quad (13)$$

Differentiating equation (13) with respect to t yields

$$\dot{x}(t) = \dot{A}\cos\phi - A\sin\phi(\omega - \dot{\psi}) + \omega A_0\epsilon H(\phi - \phi^*). \quad (14)$$

To achieve the desired form of velocity given by equation (7), the following condition is imposed:

$$\dot{A}\cos\phi + A\dot{\psi}\sin\phi = 0. \quad (15)$$

Differentiating equation (14) and taking into account equation (15) yields the expression for acceleration

$$\ddot{x}(t) = -\omega\dot{A}\sin\phi - \omega A\cos\phi(\omega - \dot{\psi}) + 2A_0\epsilon\omega^2 J(\phi - \phi^*) \quad (16)$$

in which $J(\phi - \phi^*)$ denotes a sequence of delta functions occurring periodically with frequency 2ω and alternating in sign. Substituting equations (13) and (16) into equation (3) results in

$$(\omega_n^2 - \omega^2)A\cos\phi - \omega\dot{A}\sin\phi + \omega A\dot{\psi}\cos\phi + \omega_n^2 A_0\epsilon G(\phi - \phi^*) + 2A_0\epsilon\omega^2 J(\phi - \phi^*) = \mu f(A, \phi) + K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*]) + B\cos\omega t. \quad (17)$$

Note that the impulsive forcing term, $I(x, \dot{x}^-)|\dot{x}|\delta(x)$ in equation (3) has been written as $K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*])$, consistent with equation (13).

Equations (15) and (17) are two equations in $\dot{A}$ and $\dot{\psi}$ which may be solved to yield

$$\dot{A} = \frac{1}{\omega}\{(\omega_n^2 - \omega^2)A\cos\phi\sin\phi + \omega_n^2 A_0\epsilon G(\phi - \phi^*)\sin\phi - \mu f(A, \phi)\sin\phi - B\cos(\phi + \psi)\sin\phi + (2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*]))\sin\phi\} \quad (18)$$

$$\dot{\psi} = \frac{1}{A\omega}\{-(\omega_n^2 - \omega^2)A\cos^2\phi - \omega_n^2 A_0\epsilon G(\phi - \phi^*)\cos\phi + \mu f(A, \phi)\cos\phi + B\cos(\phi + \psi)\cos\phi - (2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*]))\cos\phi\}. \quad (19)$$

Assuming A and ψ are slowly varying functions of time, $\dot{A}$ and $\dot{\psi}$ are replaced by their average values over one period

$$2\omega\dot{A} = \frac{1}{\pi}\int_0^{2\pi}\omega_n^2 A_0\epsilon G(\phi - \phi^*)\sin\phi d\phi - \frac{\mu}{\pi}\int_0^{2\pi} f(A, \phi)\sin\phi d\phi + B\sin\psi + \frac{1}{\pi}\int_0^{2\pi}(2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*]))\sin\phi d\phi \quad (20)$$

$$2\omega A\dot{\psi} = -(\omega_n^2 - \omega^2)A - \frac{1}{\pi}\int_0^{2\pi}\omega_n^2 A_0\epsilon G(\phi - \phi^*)\cos\phi d\phi + \frac{\mu}{\pi}\int_0^{2\pi} f(A, \phi)\cos\phi d\phi + B\cos\psi - \frac{1}{\pi}\int_0^{2\pi}(2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*]))\cos\phi d\phi. \quad (21)$$

The integral expressions in equations (20) and (21) containing $G(\phi - \phi^*)$ can be evaluated as follows:

$$\frac{1}{\pi}\int_0^{2\pi}\omega_n^2 A_0\epsilon G(\phi - \phi^*)\sin\phi d\phi = -\frac{4}{\pi}\omega_n^2 A_0\epsilon\cos\phi^* \quad (22)$$

$$\frac{1}{\pi}\int_0^{2\pi}\omega_n^2 A_0\epsilon G(\phi - \phi^*)\cos\phi d\phi = +\frac{4}{\pi}\omega_n^2 A_0\epsilon\sin\phi^*. \quad (23)$$

Equations (20) and (21) can then be written as

$$2\omega\dot{A} = -\frac{4}{\pi}\omega_n^2 A_0\epsilon\cos\phi^* - S(A) + \tilde{S}(A) + B\sin\psi \quad (24)$$

$$2\omega A\dot{\psi} = -(\omega_n^2 - \omega^2)A - \frac{4}{\pi}\omega_n^2 A_0\epsilon\sin\phi^* + C(A) - \tilde{C}(A) + B\cos\psi \quad (25)$$

in which

$$S(A) = \frac{\mu}{\pi}\int_0^{2\pi} f(A, \phi)\sin\phi d\phi \quad (26)$$

$$C(A) = \frac{\mu}{\pi}\int_0^{2\pi} f(A, \phi)\cos\phi d\phi \quad (27)$$

$$\tilde{S}(A) = \frac{1}{\pi}\int_0^{2\pi}\{2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*])\}\sin\phi d\phi \quad (28)$$

$$\tilde{C}(A) = \frac{1}{\pi}\int_0^{2\pi}\{2A_0\epsilon\omega^2 J(\phi - \phi^*) - K(A, \phi)\delta(A\cos\phi + A_0\epsilon G[\phi - \phi^*])\}\cos\phi d\phi. \quad (29)$$

To determine the steady-state amplitude and phase, $\dot{A}$ and $\dot{\psi}$ are set equal to zero. It may be shown (Appendix) that under steady-state conditions $\tilde{S}(A)$ and $\tilde{C}(A)$ vanish, leaving the reduced system of equations

$$\frac{4}{\pi}\omega_n^2 A_0\epsilon\cos\gamma + S(A_0) = B\sin\psi_0 \quad (30)$$

$$(\omega^2 - \omega_n^2)A_0 + \frac{4}{\pi}\omega_n^2 A_0\epsilon\sin\gamma + C(A_0) = -B\cos\psi_0. \quad (31)$$

Comparing equations (30) and (31) with those obtained using the standard SVP solution equation (4), only minor modifications are evident—namely, the addition of one new term in each equation containing ϵ and γ.

Eliminating the phase ψ_0 from the previous two equations yields an expression for the steady-state amplitude A_0 as a function of frequency and amplitude of excitation

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\left[\frac{4}{\pi}\omega_n^2 A_0\epsilon\cos\gamma + S(A_0)\right]^2$$

$$+\left[(\omega^2-\omega_n^2)A_0+\frac{4}{\pi}\omega_n^2 A_0\epsilon\sin\gamma + C(A_0)\right]^2 = B^2 \quad (32)$$

and the steady-state phase ψ_0 is obtained from

$$\tan\psi_0 = \frac{-[4/\pi\ \omega_n^2 A_0\epsilon\cos\gamma + S(A_0)]}{[(\omega^2-\omega_n^2)A_0+4/\pi\ \omega_n^2 A_0\epsilon\sin\gamma + C(A_0)]} \quad (33)$$

Example 1: Single Degree-of-Freedom Oscillator

The response of a single degree-of-freedom oscillator subject to harmonic excitation and prescribed impulses is investigated. It should be noted that due to the nature of the impulsive forces, the system under consideration is nonlinear. A piecewise-exact solution is obtained for the steady-state response, along with an approximate solution using the modified slowly varying parameter approach. In that a piecewise-exact solution to the problems exists, it appears somewhat inconsistent to apply an SVP solution technique. However, this example serves two purposes. First, it demonstrates that for a range of frequencies, the piecewise-exact solution can be expressed as the sum of two functions which are approximated well by the modified steady-state solution presented earlier; i.e., equation (12) is indeed an appropriate solution form for equation (3). Second, the accuracy of the modified solution is verified using a simple problem with known solution.

The equation of motion to be investigated is given by

$$\ddot{x} + \omega_n^2 x = B\cos\omega t. \quad (34)$$

The impulses applied to the system are of constant specified magnitude, in the direction of motion, i.e.,

$$\dot{x}^- + \mathrm{sgn}(\dot{x}^-)I_0 = \dot{x}^+ \text{ at } x = 0, \quad (35)$$

in which I_0 is the magnitude of impulse per unit mass.

Equations (34) and (35) are nondimensionalized by letting $\tau = \omega_n t$, such that

$$\frac{d}{dt} = \omega_n\frac{d}{d\tau} = \omega_n(\,)' \quad (36)$$

$$\frac{d^2}{dt^2} = \omega_n^2\frac{d^2}{d\tau^2} = \omega_n^2(\,)'' \quad (37)$$

resulting in

$$X'' + X = \cos\beta\tau \quad (38)$$

$$X'^- + \mathrm{sgn}(X'^-)\delta = X'^+ \text{ at } X = 0 \quad (39)$$

in which $\delta = I_0\omega_n/B$, $\beta = \omega/\omega_n$, and $X = x\omega_n^2/B$.

The conditions for a symmetric steady-state response at the excitation frequency are

$$\left.\begin{array}{ll} X(0) = 0 & X(\pi/\beta) = 0 \\ X'(0) = X_s' & X'(\pi/\beta) = -X_s' + \delta \end{array}\right\} \quad (40)$$

in the domain $0 \le \tau \le \pi/\beta$.

Piecewise Exact Solution. The general solution to equation (38) is given by

$$X(\tau) = \begin{cases} \rho\cos(\tau+\psi) + D\cos(\beta\tau+\phi); & 0 < \tau < \pi/\beta \\ -\rho\cos(\tau-\pi/\beta+\psi) + D\cos(\beta\tau+\phi); & \pi/\beta < \tau < 2\pi/\beta \end{cases} \quad (41)$$

in which $D = 1/(1-\beta^2)$, provided $\beta \neq 1/(2i-1)$; $i = 1, 2, 3, \ldots$ The first term in the solution, $\rho\cos(\tau+\psi)$, will be referred to as the transient term and the second, $D\cos(\beta\tau+\phi)$, as the periodic term.

Substituting the general solution equation (41) into equation (40) results in four equations for the four unknowns ρ, ψ, ϕ, and X_s'. Solving for these quantities

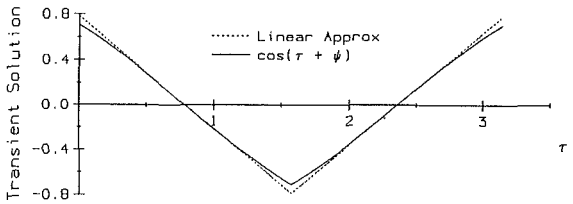


Fig. 3 Structure of solution: $\beta = 2.0$

$$\rho = -\frac{\delta}{2\cos\pi/2\beta} \quad (42)$$

$$\tan\psi = \cot\frac{\pi}{2\beta}\left(\rightarrow\psi = \frac{\pi}{2}\left[1-\frac{1}{\beta}\right]\right) \quad (43)$$

$$\cos\phi = \frac{\delta}{2D}\tan\frac{\pi}{2\beta};\ \frac{\pi}{2} \le \phi \le \frac{3\pi}{2};\ \delta > 0 \quad (44)$$

$$X_s' = \frac{\delta}{2} \pm \beta D\left\{1-\left(\frac{\delta}{2D}\tan\frac{\pi}{2\beta}\right)^2\right\}^{1/2};\ \delta < \frac{2D}{\tan\frac{\pi}{2\beta}}. \quad (45)$$

The peak amplitude of response occurs when the velocity is zero, i.e.,

$$X'(\tau_{\max}) = -\rho\sin(\tau_{\max}+\psi) - \beta D(\beta\tau_{\max}+\phi) = 0. \quad (46)$$

Equation (46) must be solved numerically for $\tau_{\max}$ and the result substituted into equation (41) to obtain the peak amplitude.

Structure of the Piecewise-Exact Solution. It is valuable to investigate the structure of the solution presented and compare it to the modified form of solution proposed in equation (6). For this, consider the characteristics of the solution equation (41) over three frequency ranges: $\beta > 2$, $0.5 < \beta < 2$ and $\beta < 0.5$.

Case (i): $\beta > 2$

The contribution to the total solution from the transient terms $\rho\cos(\tau+\psi)$ and $\rho\cos(\tau-\pi/\beta+\psi)$ for a representative $\beta = 2$ are sketched in Fig. 3. Expanding $\cos(\tau+\psi)$ in a Taylor series about $\tau = \pi/2\beta$, using equation (43), yields

$$\cos(\tau+\psi) = -(\tau-\pi/2\beta) + \frac{1}{6}(\tau-\pi/2\beta)^3 + o(\tau-\pi/2\beta)^5. \quad (47)$$

It can be seen that, in this frequency range the previous expression can be approximated by $-(\tau-\pi/2\beta)$ with reasonable accuracy (recall $0 \le \tau \le \pi/\beta$), as shown by the dashed line in Fig. 3. The maximum relative error between the exact and linear approximation is on the order of ten percent for $\beta = 2$; the approximation improves with increasing β. A similar result is obtained by expanding $\cos(\tau-\pi/\beta+\psi)$ about $3\pi/2\beta$.

This piecewise-linear approximation to the transient terms of the exact solution is the function G, defined in equation (8). The periodic terms in equation (41) are present (in exact form) in the modified solution, equation (6). In this case the modified solution is a reasonable approximation to the exact solution, equation (41).

Case (ii). $0.5 < \beta < 2$ $(\beta \neq 1)$.

In this case, it is convenient to express the transient terms as follows:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

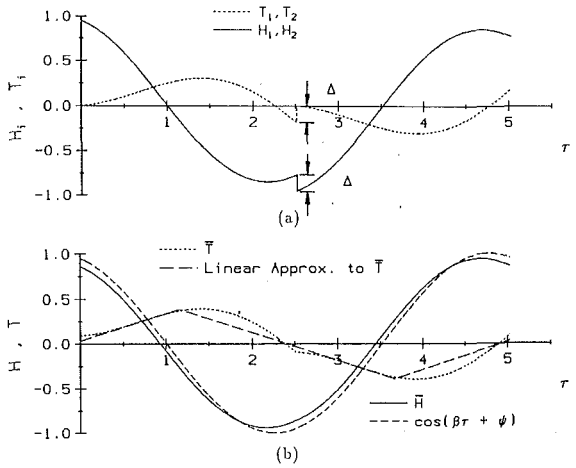


Fig. 4 Structure of solution: $\beta = 1.25$

$$\cos(\tau + \psi) = \cos[(1 - \beta)\tau + \beta\tau + \psi]$$
$$= \overbrace{\{\cos[(1 - \beta)\tau]\cos(\beta\tau + \psi)}^{H_1}$$
$$\overbrace{- \sin[(1 - \beta)\tau]\sin(\beta\tau + \psi)\}}^{T_1} \quad (48)$$

$$-\cos(\tau - \pi/\beta + \psi) = -\cos[(1 - \beta)\tau - \pi/\beta + \beta\tau + \psi]$$
$$= \underbrace{\{-\cos[(1 - \beta)\tau - \pi/\beta]\cos(\beta\tau + \psi)}_{H_2}$$
$$\underbrace{+ \sin[(1 - \beta)\tau - \pi/\beta]\sin(\beta\tau + \psi)\}}_{T_2}. \quad (49)$$

The functions H_1, H_2, T_1, T_2 indicated above are plotted for a representative $\beta = 1.25$ in Fig. 4(*a*). Note, the discontinuities between the H and T functions at π/β, denoted by Δ, are equal and given by

$$\Delta = \sin\frac{\pi}{\beta}\cos\frac{\pi}{2\beta}. \quad (50)$$

Defining new functions $\bar{H}$ and $\bar{T}$ as (Fig. 4(*b*))

$$\bar{H} = \begin{cases} H_1 - \Delta/2 \\ H_2 + \Delta/2 \end{cases}, \quad (51)$$

$$\bar{T} = \begin{cases} T_1 + \Delta/2 \\ T_2 - \Delta/2 \end{cases} \quad (52)$$

the transient solution may be represented as the sum $\bar{H}$ and $\bar{T}$ over the full period.

The shifted functions are now continuous and periodic with frequency β. For $\beta \approx 1$, $\cos[(1-\beta)\tau]$ and $\cos[(1-\beta)\tau - \pi/\beta]$ represent slowly varying amplitudes. The function $\bar{H}$ can be therefore approximated by $\cos(\beta\tau + \psi)$, shown by the dashed line in Fig. 4(*b*). Similarly, $\bar{T}$ can be approximated by a linear triangular wave (e.g., the function G, defined in equation (8)), also shown by a dashed line in Fig. 4(*b*). As β tends to 1, the approximation improves. In this frequency range, equation (6) is still a good approximation to the exact solution equation (41).

Case (iii). $\beta < 0.5$.

Below $\beta = 0.5$, the piecewise-exact solution can no longer be approximated by equation (6). In this range, the periodic terms of equation (41) are present without approximation in the modified solution equation (6). However, the transient terms now oscillate at a frequency $\geq 2\beta$, and are therefore not approximated well by the linear function G. Use of this solution technique is not appropriate in this frequency range.

It is noted that the structure of solution in this range of frequencies is very rich, exhibiting subharmonic responses and solutions periodic at the excitation frequency for only restricted values of $\delta = \delta(\beta)$. Detailed discussion of this structure is beyond the scope of this paper.

Modified SVP Solution. In dimensionless form, the solution to equations (38)–(39) may be expressed as

$$X(\tau) = A(\tau)\cos[\beta\tau - \psi(\tau)] + A_0\epsilon G(\beta\tau - \psi_0 + \gamma). \quad (53)$$

Once again, using the notation $A(\tau) = A$, $\psi = \psi(\tau)$, $\phi = \beta\tau - \psi(\tau)$, $\phi^* = -\psi + \psi_0 - \gamma$, this can be written as

$$X(\tau) = A\cos\phi + A_0\epsilon G(\phi - \phi^*). \quad (54)$$

The impulse-momentum condition equation (39) is satisfied during steady-state by solution of the following equations

$$X(\tau)|_{\phi_1} = 0 \quad (55)$$

$$X'(\tau)|_{\phi_1^-} + \text{sgn}(X')\delta|_{\phi_1^-} = X'(\tau)|_{\phi_1^+} \quad (56)$$

in which $\phi_1 = \beta\tau_1 - \psi_0$. Solving the previous two equations for ϵ and γ yields

$$\epsilon = -\frac{\delta}{2A_0\beta} \quad (57)$$

$$\gamma = \sin^{-1}\left\{-\frac{\delta\pi}{4A_0\beta}\right\}. \quad (58)$$

Since $\mu = 0$, $S(A) = C(A) = 0$ and the equation for determining the steady-state amplitude A_0 reduces to

$$\left[\frac{4}{\pi}A_0\epsilon\cos\gamma\right]^2 + \left[(\beta^2 - 1)A_0 + \frac{4}{\pi}A_0\epsilon\sin\gamma\right]^2 = 1. \quad (59)$$

Substituting in equation (59) the expressions for ϵ and γ, and solving for A_0 yields

$$A_0 = \frac{1}{1-\beta^2}\left[1 - \frac{\delta^2}{\beta^2}(\beta^2 - 1 + 4/\pi^2)\right]^{1/2}. \quad (60)$$

The phase ψ_0 is given by

$$\tan\psi_0 = \frac{-4/\pi\,\epsilon\cos\gamma}{(\beta^2 - 1) + 4/\pi\,\epsilon\sin\gamma}. \quad (61)$$

Finally, the approximate steady-state solution is

$$X(\tau) = \frac{1}{1-\beta^2}\left[1 - \frac{\delta^2}{\beta^2}(\beta^2 - 1 + 4/\pi^2)\right]^{1/2}\{\cos(\beta\tau - \psi_0) + \epsilon G(\beta\tau - \psi_0 + \gamma)\}. \quad (62)$$

The peak response in the range $\phi_1 - \pi < \phi < \phi_1$ is

$$X_{\max} = A_0\{\cos\phi_{\max} + \epsilon G(\phi_{\max} + \gamma)\} \quad (63)$$

in which

$$\phi_{\max} = \sin^{-1}[\epsilon]. \quad (64)$$

The total phase is given by

$$\psi_t = \psi_0 - \gamma. \quad (65)$$

The peak displacement obtained from the exact equation (41) and approximate equation (62) solutions are presented in Fig. 5(*a*) as a function of β, for various δ. The results are in good agreement over the frequency range $0.5 < \beta < 2.0$. Slight deviations from the exact results occur near $\beta = 0.5$, for reasons discussed earlier. The corresponding phases are shown in Fig. 5(*b*). Once again, the results agree extremely well, with only slight deviations near $\beta = 0.5$. In each case, the error in the approximate solution increases with increasing magnitude of impulse (δ). This would be expected, as the transient effect of the impulse on the system becomes more pronounced with increasing magnitude of the impulse.

Displacement and velocity profiles are presented in Fig. 6

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

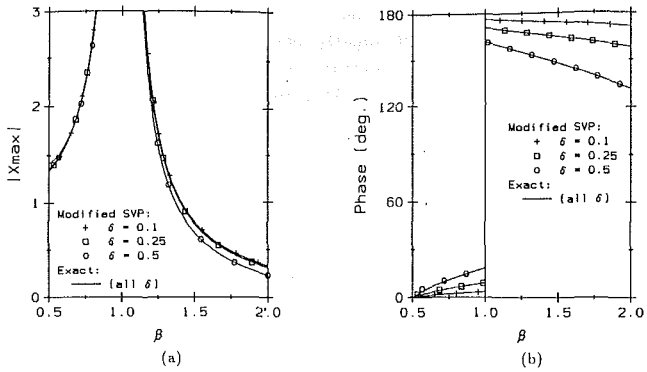


Fig. 5 Amplitude and phase ($3\pi/2 - \phi$) response—linear system with impulses

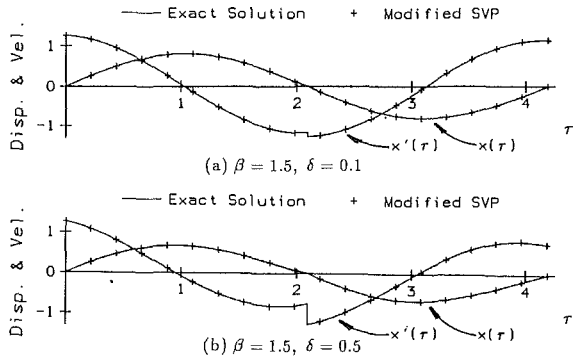


Fig. 6 Displacement and velocity profiles—linear system with impulses

for a representative $\beta = 1.5$ and various δ. The agreement with the exact solution is good over the entire period. The approximate solution accurately characterizes the true response not only in form, but in peak response and discontinuity in velocity.

Example 2: Rigid Block Subject to Harmonic Ground Motion

Provided there is sufficient friction to prevent slippage, a pure rocking response of a rigid block subjected to harmonic base motion can be observed. A number of investigators have studied this problem and recently a piecewise-exact solution was obtained by Spanos and Koh (1984).

The equation of motion for the system, shown in Fig. 7, for small angles of rotation, can be expressed as (Spanos and Koh, 1984):

$$(I + mR^2)\ddot{\theta} + mg\,R\,\text{sgn}(\theta)(\theta_c - |\theta|) = mR\ddot{x}_g \qquad (66)$$

in which m is the mass of the block, I the mass moment of inertia about the center of gravity, R the distance from the center of gravity to the corner in contact, θ_c the angle between the side of the block and the line forming R, and $\ddot{x}_g$ the horizontal ground acceleration. The signum function is defined as

$$\text{sgn}(\theta) = \begin{cases} +1 \text{ for } \theta > 0 \text{ a rotation about } O^+ \\ -1 \text{ for } \theta < 0 \text{ a rotation about } O^- \end{cases} . \qquad (67)$$

When $\theta = 0$, an impact occurs between one corner of the block and the ground with a magnitude such that the angular velocity immediately after impact is a fraction of the velocity immediately before, defined by the relation

$$\dot{\theta}^+ = \delta\dot{\theta}^- \qquad (68)$$

in which δ is derived from conservation of angular momentum before and after the impact, and is given by[3]

$$\delta = 1 - \frac{3}{2}\sin^2\theta_c. \qquad (69)$$

For a rectangular block of dimension $2h$ by $2b$, and a harmonic base motion $\ddot{x}_g = gA_g \cos \omega t$, equation (66) reduces to

$$\ddot{\theta} + \frac{\alpha^2}{4}\,\text{sgn}(\theta)(\theta_c - |\theta|) = \frac{\alpha^2}{4}A_g \cos \omega t \qquad (70)$$

in which $\alpha^2 = mgR/I$. Equation (70) is nondimensionalized by letting $\tau = \alpha t$, $\beta = \omega/\alpha$ and $\Theta = \theta/\theta_c$, i.e.,

$$\Theta'' + \frac{1}{4}\,\text{sgn}(\Theta)(1 - |\Theta|) = B \cos \beta\tau \qquad (71)$$

[3]It is assumed that a *point* contact occurs at the moment of impact. The coefficient of restitution between the corner of the block in contact and the base is taken as zero.

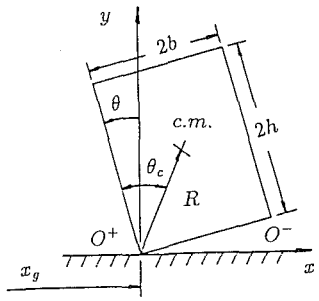


Fig. 7 Rocking rigid block—coordinate definition

in which $(\,)' = d/d\tau$ and $B = A_g/4\theta_c$. The impact condition, equation (68), becomes

$$\Theta'^+ = \delta\Theta'^-. \qquad (72)$$

Standard SVP Solution. Using the nondimensional form of equation (3), the equation of motion may be written:

$$\Theta'' + \beta_n^2\Theta = \beta_n^2\Theta - \frac{1}{4}\,\text{sgn}(\Theta)(1 - |\Theta|)$$

$$- \frac{1}{4}(1-\delta)\Theta'|\Theta'|\delta(\Theta) + B \cos \beta\tau \qquad (73)$$

where $\beta_n = \omega_n/\alpha$, and the property of the delta function, $\delta(cx) = \delta(x)/|c|$ for some constant, c, has been used. Note that to force equation (73) into the form of equation (3), the $\beta_n^2\Theta$ term appears on both sides of the equation.

Setting $\epsilon = 0$ in equations (30) and (31) yields the standard SVP equations, and noting that the $\beta_n^2\Theta$ contributions cancel, these reduce to

$$S(A_0) = B \sin \psi_0 \qquad (74)$$

$$\beta^2 A_0 + C(A_0) = -B \cos \psi_0 \qquad (75)$$

in which

$$S(A_0) = \frac{(1-\delta)}{2\pi}A_0\beta^2 \qquad (76)$$

$$C(A_0) = \frac{A_0}{4} - \frac{1}{\pi}. \qquad (77)$$

Solving for the steady-state amplitude and phase yields

$$A_0 = \frac{4/\pi \pm \{(A_g/\theta_c)^2(1+\Delta) - (4/\pi)^2\Delta\}^{1/2}}{(1+4\beta^2)(1+\Delta)} \qquad (78)$$

$$\tan \psi_0 = \frac{-2/\pi\,(1-\delta)A_0\beta^2}{(4\beta^2+1)A_0 - 4/\pi} \qquad (79)$$

in which

$$\Delta = \frac{4(1-\delta)^2}{\pi^2}\beta^4(1 + 4\beta^2)^{-2}. \qquad (80)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

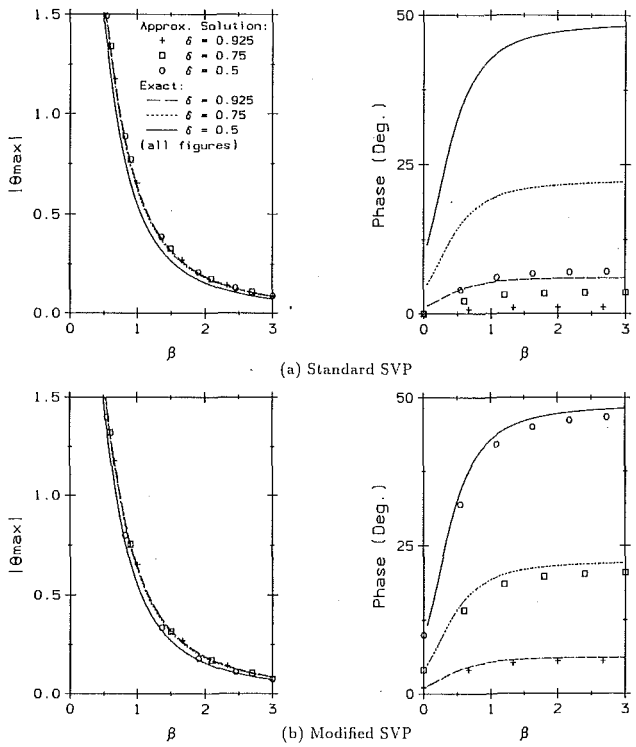


Fig. 8 Amplitude and phase response—rocking rigid block

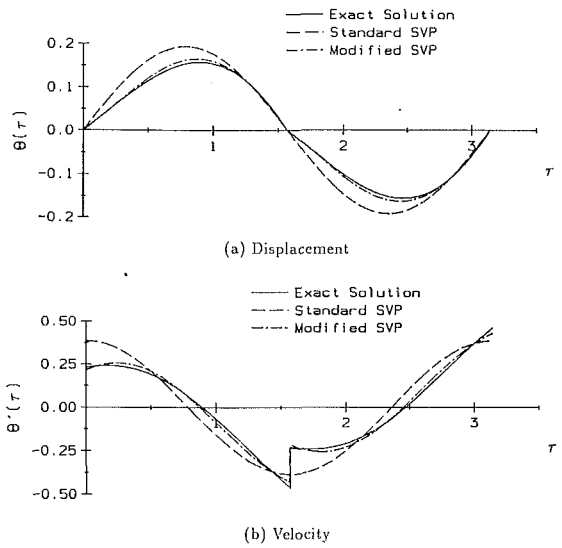


Fig. 9 Velocity and displacement profiles—rocking rigid block (β = 2.0, δ = 0.5)

Modified SVP Solution. Using the notation outlined earlier, the approximate solution in dimensionless form is expressed as

$$\Theta(\tau) = A\cos\phi + A_0\epsilon G(\phi-\phi^*). \tag{81}$$

The impulse-momentum condition equation (72) is satisfied during steady-state motion by

$$\Theta(\tau)|_{\phi_1} = 0 \tag{82}$$

$$\Theta'(\tau)|_{\phi_1^+} = \delta\Theta'(\tau)|_{\phi_1^-}. \tag{83}$$

Solving for ϵ and γ yields

$$\epsilon = \left[\left(\frac{\pi}{2}\right)^2 + \lambda^2\right]^{-1/2} \tag{84}$$

$$\gamma = \tan^{-1}\left[\frac{\pi}{2\lambda}\right]^{1/2} \tag{85}$$

where $\lambda = (1+\delta)/(1-\delta)$.

Performing the indicated integrations in equations (26) and (27) and reducing to steady-state conditions, $C(A_0)$ and $S(A_0)$ are obtained as

$$S(A_0) = \frac{1}{\pi}\{\sin\gamma - A_0\epsilon\cos\gamma\} \tag{86}$$

$$C(A_0) = \frac{A_0}{4}\left\{1 - \frac{4}{\pi}\epsilon\sin\gamma\right\} - \frac{1}{\pi}\cos\gamma. \tag{87}$$

Solving for A_0 and ψ_0 from equations (32) and (33) yields

$$A_0 = \frac{4/\pi\,\lambda\epsilon \pm \{(4/\pi\,\lambda\epsilon)^2 - \Delta[(4/\pi)^2 - (A_g/\theta_c)^2]\}^{1/2}}{(1+4\beta^2)\Delta} \tag{88}$$

$$\tan\psi_0 = \frac{\epsilon/2 - (\epsilon^2\lambda/\pi)A_0}{\beta^2 A_0 + (A_0/4)(1-2\epsilon^2) - \epsilon\lambda/\pi} \tag{89}$$

in which

$$\Delta = 1 - \frac{4\epsilon^2}{(1+4\beta^2)} + \frac{(4/\pi\,\epsilon)^2}{(1+4\beta^2)^2}. \tag{90}$$

Frequency response curves, obtained from the exact solution (Spanos and Koh, 1984) and equations (78) and (81) are shown in Fig. 8. All the results presented correspond to $A_g/\theta_c = 2.0$. Note that the figures show only solutions known to be stable (Spanos and Koh, 1984). A family of unstable solutions for the examples shown lies to the left of the frequency response curves presented. These are omitted for clarity.

The peak angle of rotation predicted by equation (78) for $\delta = .925$ is fairly accurate. This would be expected for impulses of low magnitude. However, as δ decreases the solution fails to accurately predict the peak amplitude, which is overestimated for all frequencies. The phase angle given by equation (79) and shown in Fig. 8(*a*), is poor for all values of δ and for all frequencies.

The modified SVP solution given by equation (81) compares well with the exact solution for both amplitude and phase, as shown in Fig. 8(*b*). The maximum angle of rotation is in good agreement with the exact results for all δ and over all frequencies. Likewise, the phase angle predicted shows similar agreement.

Displacement and velocity profiles for the block are presented in Fig. 9 as obtained by the three solutions. These correspond to an $A_g/\theta_c = 2.0$, $\delta = 0.5$ and $\beta = 2.0$. Note that the modified SVP solution compares well with the exact solution over the entire period. At $\theta = 0$, the change in slope of the displacement curve and discontinuity in velocity are present. In contrast, the solution given by equation (78) is a fair representation of the displacement profile, but is a poor estimator of velocity.

Conclusions

A new form of solution has been proposed for obtaining approximate periodic solutions to nonlinear systems with prescribed impulses. The assumed form has been developed for use with the method of slowly varying parameters to determine the approximate amplitude and phase of the steady-state motion. The resulting equations are only slight modifications of those obtained from the standard SVP analysis.

The modified first approximation shows marked improvements over the standard form. Under steady-state motion, the solution displays the characteristic properties of systems of this type. These include a change in slope of the displacement and discontinuity in velocity at the point of application of the impulse. These are absent from the standard form of solution which is a simple harmonic.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The results obtained from the modified solution have been compared with those from the piecewise-exact solution for two examples. In both cases, the results are in good agreement for a wide range of impulse magnitudes and frequencies. In the rocking block example, the exact results have also been compared with the approximate solution obtained using a standard slowly varying parameter approach (which yields the same results as equivalent linearization) as proposed by other researchers for problems of this type. The modified solution is demonstrated as significantly more accurate than the standard technique in predicting peak responses and in the actual representation of the response.

The proposed form of solution, while presented here for two examples where piecewise-exact solutions are available, has much broader application. For many nonlinear systems with impulses, the piecewise-exact solutions are not available. In addition, which is of significance to the rocking block problem when sliding is permitted, the method enables efficient and accurate solutions of multiple degree-of-freedom, nonlinear systems which are otherwise attainable using only a standard slowly varying parameter approach.

The technique can be expanded to include nonzero x in $I(x, \dot{x})$, and to consider subharmonic and superharmonic response, just as standard SVP. Presented herein was an introduction to the method; it has many possibilities. A significant advantage of the technique is that frequency response curves are obtained explicitly: There is no need for numerical solution of algebraic equations to obtain peak amplitudes, and other quantities of interest.

References

Bogoliubov, N. N., and Mitropolsky, Y. A., 1961, *Asymptotic Methods in the Theory of Non-Linear Oscillations*, Hindvstan Publishing Corp., India.

Butenin, N. V., 1965, *Elements of the Theory of Nonlinear Oscillations*, Blaisdell Publishing Co., New York.

Caughey, T. K., 1963, "Equivalent Linearization Techniques," *J. Acoust. Soc. Am.*, Vol. 35, pp. 1706–1711.

Iwan, W. D., 1969, "On Defining Equivalent Systems for Certain Ordinary Non-Linear Differential Equations," *Int. J. Non-Linear Mech.*, Vol. 4, pp. 325–334.

Krylov, N. N., and Bogoliubov, N. N., 1970, *Introduction to Non-Linear Mechanics*, Princeton University Press, N.J. 1947, Kraus Reprint Company, N.Y.

Masri, S. F., and Caughey, T. K., 1966, "On the Stability of the Impact Damper," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 586–592.

Mitropolsky, Y. A., and Samoilenko, A. M., 1985, "Forced Oscillations of Systems with Impulsive Force," *Int. J. Non-Linear Mech.*, Vol. 20, No. 5–6, pp. 419–426.

Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York.

Samoilenko, A. M., 1961, "Application of the Averaging Method for Investigation of Oscillations Generated by Transient Impulses in Auto-Oscillations in Systems of Second Order with a Small Parameter," *Ukr. Mat. Zh.*, Vol. 13, No. 3.

Samoilenko, A. M., 1965, "On the Investigation of Differential Equations with an Irregular Right-Hand Side," *Third Conf. on Non-Linear Oscillations*, Vol. 1, Akademie-Verlag, Berlin, pp. 106–113.

Samoilenko, A. M., 1967, "Averaging Method for Investigating Systems Subjected to an Impulsive Action," *Ukr. Mat. Zh.*, Vol. 19, No. 5, pp. 96–104.

Samoilenko, A. M., and Perestyuk, N. A., 1974, "The Method of Averaging in Systems with an Impulsive Action," Translated from *Ukr. Mat. Zh.*, Vol. 26, No. 3, pp. 411–418.

Spanos, P. D., and Koh, A., 1984, "Rocking of Rigid Blocks Due to Harmonic Shaking," *J. Eng. Mech.*, Vol. 110, No. 11, ASCE, New York.

APPENDIX

The derivation presented herein verifies that $\tilde{S}(A)$ and $\tilde{C}(A)$ given in equations (28) and (29) vanish in the steady-state condition.

Consider the integration of the first term in the integrand of equation (28):

$$\int_0^{2\pi} 2A_0\epsilon\omega^2 J(\phi - \phi^*) \sin \phi d\phi \tag{A1}$$

in which $J(\phi - \phi^*)$ denotes a sequence of delta functions occurring periodically with frequency 2ω and alternating in sign, i.e.,

$$J(\phi - \phi^*) = \delta(\phi - \phi^* - \pi/2) - \delta(\phi - \phi^* - 3\pi/2) \tag{A2}$$

for $\phi^* \le \phi \le \phi^* + 2\pi$. Equation (A2) is consistent with the definitions of G and H given in equations (8) and (9). Under steady-state conditions, equations (A1) and (A2) reduce to

$$\int_0^{2\pi} 2A_0\epsilon\omega^2 J(\phi_0 + \gamma) \sin \phi_0 d\phi_0 \tag{A3}$$

and

$$J(\phi_0 + \gamma) = \delta(\phi_0 - [\pi/2 - \gamma]) - \delta(\phi_0 - [3\pi/2 - \gamma]) \tag{A4}$$

for $\gamma \le \phi_0 \le \gamma + 2\pi$. Substituting equation (A4) into (A3) and performing the indicated integration yields

$$\int_0^{2\pi} 2A_0\epsilon\omega^2 J(\phi_0 + \gamma) \sin \phi_0 d\phi_0 = 4A_0\epsilon\omega^2 \cos \gamma. \tag{A5}$$

Consider now the integration of the second term in equation (28), i.e.,

$$\int_0^{2\pi} K(A, \phi)\delta(A \cos \phi + A_0\epsilon G(\phi - \phi^*)) \sin \phi d\phi. \tag{A6}$$

In terms of x and $\dot{x}$, the above integral can be expressed as

$$\int_0^{2\pi} I(x, \dot{x}^-)|\dot{x}|\delta(x) \sin \phi d\phi \tag{A7}$$

in which

$$\dot{x}^- + I(x, \dot{x}^-) = \dot{x}^+. \tag{A8}$$

Under steady-state conditions, equation (A7) reduces to

$$\int_0^{2\pi} I_0|\dot{x}|\delta(x) \sin \phi_0 d\phi_0 \tag{A9}$$

in which I_0 denotes the steady-state impulse per unit mass, given by

$$\begin{aligned} I_0 = \dot{x}^+ - \dot{x}^- &= [-\omega A_0 \sin \phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]^+ \\ &\quad - [-\omega A_0 \sin \phi_0 + \omega A_0\epsilon H(\phi_0 + \gamma)]^- \\ &= \omega A_0\epsilon [H^+(\phi_0 + \gamma) - H^-(\phi_0 + \gamma)]. \end{aligned} \tag{A10}$$

With the change of variables,

$$\phi_0 = \omega t - \psi_0;\ d\phi_0 = \omega dt = \omega dx/\dot{x}, \tag{A11}$$

equation (A9) can be written as

$$-\int_{x=0^+}^{x=0^-} \omega \sin \phi_0 I_0\delta(x)dx + \int_{x=0^-}^{x=0^+} \omega \sin \phi_0 I_0\delta(x)dx. \tag{A12}$$

The first integral above corresponds to the impulse at $\phi_0 = \pi/2 - \gamma$, assuming $\dot{x} < 0$ when $x = 0$. Likewise, the second integral corresponds to the impulse at $\phi_0 = 3\pi/2 - \gamma$, assuming $\dot{x} > 0$ when $x = 0$. Evaluating equation (A12) yields

$$[\omega \sin \phi_0 I_0]_{\phi_0 = \pi/2 - \gamma} + [\omega \sin \phi_0 I_0]_{\phi_0 = 3\pi/2 - \gamma}. \tag{A13}$$

Noting that from equation (A10),

$$I_0|_{\phi_0 = \pi/2 - \gamma} = 2\omega A_0\epsilon;\quad I_0|_{\phi_0 = 3\pi/2 - \gamma} = -2\omega A_0\epsilon, \tag{A14}$$

equation (A13) reduces to

$$\omega \cos \gamma(2\omega A_0\epsilon) - \omega \cos \gamma(-2\omega A_0\epsilon) = 4A_0\epsilon\omega^2 \cos \gamma. \tag{A15}$$

The results of the two integrations (equations (A5) and (A15)) are identical and therefore equation (28) vanishes under steady-state conditions. A similar development follows for equation (29).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

N. Sri Namachchivaya
Aeronautical and Astronautical Engineering Department, University of Illinois at Urbana-Champaign, Urbana, IL 61801-2997
Mem. ASME

Co-dimension Two Bifurcations in the Presence of Noise

Some results pertaining to co-dimension two stochastic bifurcations are presented. The normal form associated with non-semi-simple double-zero eigenvalues is considered. The method of stochastic averaging applicable for singularly perturbed stochastic differential equations is used to further reduce the problem to a one-dimensional diffusion process. Probability density, most probable values, stability conditions in probability, and mean exit times are obtained for the reduced system.

Introduction

There exist several natural phenomena that vary in a random manner due to the effects of large numbers of unknown factors. Dynamic systems in such environments are subjected to stochastic excitations. The stability and the nonlinear response of such stochastic systems have become of increasing interest in engineering. Examples of stochastic excitations include forces generated by jet and rocket engines in modern high-powered aircraft, space and missile structures, as well as excitation due to earthquakes, ocean waves, and wind gusts. They fluctuate randomly over a wide band of frequencies and have to be considered as stochastic functions of time defined only in probabilistic terms. The effects of such fluctuations on nonlinear systems have been studied by various researchers. For one-dimensional systems, Horsthemke and Lefever (1977) and Arnold et al. (1978) established the existence of transition or bifurcations solely induced by noise. These ideas were extended to two-dimensional systems by Arnold et al. (1979) using the well-known Lotka-Voltera model for two interacting populations. The effect of stochastic perturbations upon a dynamical system exhibiting co-dimension one bifurcations has been studied by Baras et al. (1982), Graham (1982), Lefever and Turner (1984), and the author (1988a, 1988b, 1989). The effect of additive noise was considered in (1982), while the effect of multiplicative noise was studied by Graham (1982). The author (1988b, 1989) considered a more general problem in R^n, with both multiplicative and additive stochastic excitations. Even though a large amount of work in co-dimension one stochastic bifurcations has been reported, there is still considerably more work to be done in stochastically perturbed, co-dimension two bifurcations. The goal of this paper is to present the results pertaining to the statistical, as well as the sample, behavior of nonlinear systems that exhibit a particular co-dimension two bifurcation and are subjected to random excitations.

A dynamical system undergoes a co-dimension two bifurcation due to the presence of additional degeneracies other than those encountered for the simple and the Hopf bifurcations. Such degeneracies can be classified into two types: In the first, the linear part is similar to that of co-dimension one bifurcation but additional degeneracies occur in the nonlinear terms of the normal form or higher-order degeneracies occur in the linear part. In the second, the linear part of the vector field is doubly degenerate. In this paper, we shall consider the stochastic version of the co-dimension two bifurcation of the latter type. There are three cases of such types of degeneracy with two, three, and four-dimensional center manifold where the linear part can have: (1) two zero eigenvalues, (2) a pair of pure imaginary eigenvalues and a zero eigenvalue, and (3) two pairs of pure imaginary eigenvalues without 1:1 resonance, respectively. For a deterministic symmetric system these are represented by the following three cases which depend on two parameters μ_1 and μ_2:

(1) Double zero, nondiagonalizable eigenvalues

$$\ddot{x}+\mu_2\dot{x}+\dot{x}x^2+\mu_1 x\pm x^3=0. \tag{1a}$$

(2) Simple zero and pure imaginary pair of eigenvalues

$$\dot{r}=\mu_1 r+arz+(cr^3+drz^2),$$
$$\dot{z}=\mu_2+br^2-z^2+(er^2z+fz^3). \tag{1b}$$

(3) Two pure imaginary pairs of eigenvalues without resonance

$$\dot{r}_1=\mu_1 r_1+a_{11}r_1{}^3+a_{12}r_1r_2{}^2,\ \dot{r}_2=\mu_2 r_2+a_{21}r_1{}^2r_2+a_{22}r_2{}^3. \tag{1c}$$

The detailed analysis of these cases with various classifications and unfoldings are summarized in Guckenheimer and Holmes (1983). For simplicity, in this paper we shall restrict our attention to the stochastic version of case 1, i.e., double-zero eigenvalues with non-semi-simple forms as in Graham (1987).

Statement of the Problem

Consider a co-dimension two bifurcation associated with non-semi-simple double-zero eigenvalues, whose center manifold is two dimensional and the associated normal form is given by Guckenheimer and Holmes (1983)

$$\dot{u}=v,$$
$$\dot{v}=\mu_1\mu+\mu_2 v\pm u^3-u^2v \tag{2}$$

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 31, 1989; final revision, Nov. 16, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

where μ_1 and μ_2 are the unfolding parameters and $\mu_1=\mu_2=0$ define the co-dimension two singularity point. We are interested here in the case where the normal form is perturbed by weak Gaussian white noise and we assume without proof that the associated normal form is obtained by letting $\mu_1=\mu_0(1+\epsilon^{1/2}\eta_1(t))$ and by introducing an additive noise term $\epsilon^{1/2}\eta_2(t)$ in the second equation. It is worth noting that the influence of additive noise for this case was also considered by Graham (1987) and, as in his case, the emphasis will be given more to the bifurcation behavior as opposed to the derivation of the stochastic normal forms. Moreover, since the normal form for this case represents the van der Pol-Duffing oscillator, the problem below, given by equation (3), can also be viewed as a van der Pol-Duffing oscillator under both parametric and external excitations. Detailed analysis of the stochastic normal forms for various nonlinear stochastic systems have been given by Sri Namachchivaya (1989). Thus, consider

$$\dot{u}=v,$$
$$\dot{v}=\mu_0 u+\mu_2 v\pm u^3-u^2v+\epsilon^{1/2}[\mu_0 u\eta_1(t)+\eta_2(t)]. \tag{3}$$

For the linear equation, since the stability is not affected by the nonhomogeneous terms, the following set of Itô equations is examined:

$$\dot{x}=Bx+\sigma x\, dw$$

where

$$x=(u,\ v),\quad B=\begin{bmatrix}0 & 1\\ \mu_0 & \mu_2\end{bmatrix},\quad \sigma=\begin{bmatrix}0 & 0\\ \nu & 0\end{bmatrix}$$

and the excitation term is $\mu_o\eta(t)dt=\nu dw$. The almost-sure asymptotic stability condition is obtained, using the approach of Khasminskii (1967), by examining the Lyapunov exponent λ. For the above equation, the value of λ can be obtained in the same manner as in Kozin and Prodromou (1971) and Nishioka (1976) and is explicitly given as

$$\lambda=\int_0^{2\pi} Q(\theta)p(\theta)d\theta$$

where

$$Q(\theta)=(\mu_0+1)\cos\theta\ \sin\theta+\mu_2\sin^2\theta+\frac{1}{2}\nu^2\cos^2\theta\ \cos 2\theta$$

$$p(\theta)=\frac{C\,S(-\pi/2,\theta]}{\nu^2\cos^4\theta\ s(\theta)} \qquad -\pi/2\leqq\theta<\pi/2$$

$$p(\theta-\pi) \qquad \pi/2\leqq\theta\leqq 3\pi/2$$

$$S(-\pi/2,\theta)=\int_{-\pi/2}^{\theta} s(\xi)d\xi,$$

$$s(\xi)=\frac{1}{\cos^2\xi}\exp\left[\frac{\tan\xi}{3\nu^2}(f(\xi))\right]$$

$$f(\xi)=2\tan^2\xi-3\mu_2\tan\xi-6\mu_o$$

and C is a normalizing constant. If λ is negative, the sample paths are stable with probability one. However, in this paper our attention will be focused on the analysis of the nonlinear system.

Analysis

It is obvious that for both $\pm u^3$, the two-dimensional system undergoes stochastically perturbed co-dimension one bifurcations when the parameters μ_0 and μ_2 take the values $\mu_2\neq 0$, $\mu_0=0$ (simple) and $\mu_0<0$ and $\mu_2=0$ (Hopf) and have been studied before (Arnold et al., 1978; Sri Namachchivaya, 1988). Thus, in this section, we consider the bifurcations associated with the reduced normal form which represents a weakly perturbed conservative system. Such a reduction is obtained by using the rescalings

$$\mu_0=\epsilon^2\nu_0,\ \mu_2=\epsilon^2\nu_2,\ u=\epsilon\bar{u},\ v=\epsilon^2\bar{v},\ \bar{t}=\epsilon t,$$

and henceforth, omitting the bars from the scaled variables for simplicity, we can reduce equation (3) to the form

$$\frac{du}{dt}=v.$$

$$\frac{dv}{dt}=\nu_0 u\pm u^3+\epsilon[(\nu_2-u^2)v]+\epsilon^{1/2}[\nu_0 u\eta_1(t)+\eta_2(t)], \tag{4}$$

which can be interpreted as either a Stratonovich or an Itô equation since the correction term is identically zero. Now introducing $H=v^2/2+P_{1,2}(u)$, $P_{1,2}(u)=-\nu_0 u^2/2\pm u^4/4$, $G(u)=(\nu_2-u^2)$, the Itô equations for u and H can be written as

$$\begin{aligned}du&=\sqrt{Q_{1,2}(u)}\,dt\\ dH&=\epsilon\{Q_{1,2}(u)G(u)+F(u)\}dt\\ &\quad+\epsilon^{1/2}\sqrt{Q_{1,2}(u)}\{\sigma_r^i(u)\ dw_i\}\end{aligned} \tag{5}$$

where

$$Q_{1,2}(u)=2(H-P_{1,2}(u)),\ F(u)=\frac{1}{2}\sum_{i=1}^{2}(\sigma_v^i(u))^2,$$

$$\sigma_v^1=\nu_0 u,\ \sigma_v^2=1.$$

In the sequel we consider only two cases, namely:

(i) $H=v^2/2+P_1(u)$, $\nu_o<0$, and the fixed points are given by (0, 0) $\leftrightarrow$ stable and $(\pm\sqrt{-\nu_0},\ 0)$ $\leftrightarrow$ unstable. The Hamiltonian levels of interest lie within $H=(0,\ \nu_0^2/4)$ and $u_0^{\pm}=\pm[-\nu_o+(\nu_o^2-4H)^{1/2}]^{1/2}$.

(ii) $H=v^2/2+P_2(u)$, $\nu_o>0$, and the fixed points are given by (0, 0) $\leftrightarrow$ unstable and $(\pm\sqrt{\nu_0},\ 0)$ $\leftrightarrow$ stable. The Hamiltonian levels of interest lie within $H=(-\nu_0^2/4,\ 0)$ and $u_0^{\pm}=[\nu_0\pm(\nu_0^2+4H)^{1/2}]^{1/2}$.

Now, applying the theorem of Khasminskii (1968), we obtain the one-dimensional Itô equation

$$dH=\bar{A}(H)\ dt+\bar{\sigma}_{HH}(H)\ dw \tag{6}$$

where

$$\bar{A}(H)=\left(\frac{1}{\Lambda(H)}\right)\int_{u_o^-(H)}^{u_o^+(H)}\left\{G(u)\sqrt{Q_{1,2}(u)}+\frac{F(u)}{\sqrt{Q_{1,2}(u)}}\right\}du=\frac{\{B(H)+C(H)\}}{\Lambda(H)}$$

$$B(H)=\left(\frac{2}{15}\right)\left(\frac{\nu_0}{2-m}\right)^{5/2}\{[5(\nu_2/\nu_o)(2-m)^2-4(m^2-m+1)]E(m)-2[5(\nu_2/\nu_o)(2-m)(1-m)-(2-m)(1-m)]F(m)\}$$

$$C(H)=\left(\frac{\nu_0}{2-m}\right)^{1/2}[(S_{11}\nu_o^2)E(m)+(S_{22}/2\nu_0)(2-m)F(m)]$$

$$\bar{\sigma}_{HH}^2(H)=\left(\frac{1}{\Lambda(H)}\right)\int_{u_o^-(H)}^{u_o^+(H)}\sqrt{Q_{1,2}(u)}\{2F(u)\}du=\frac{\sigma_{HH}^2}{\Lambda(H)},$$

$$\sigma_{HH}^2(H)=\left(\frac{2}{15}\right)\left(\frac{\nu_0}{2-m}\right)^{5/2}\{[5(S_{22}/\nu_0)(2-m)^2+4S_{11}\nu_o^2(m^2-m+1)]E(m)$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

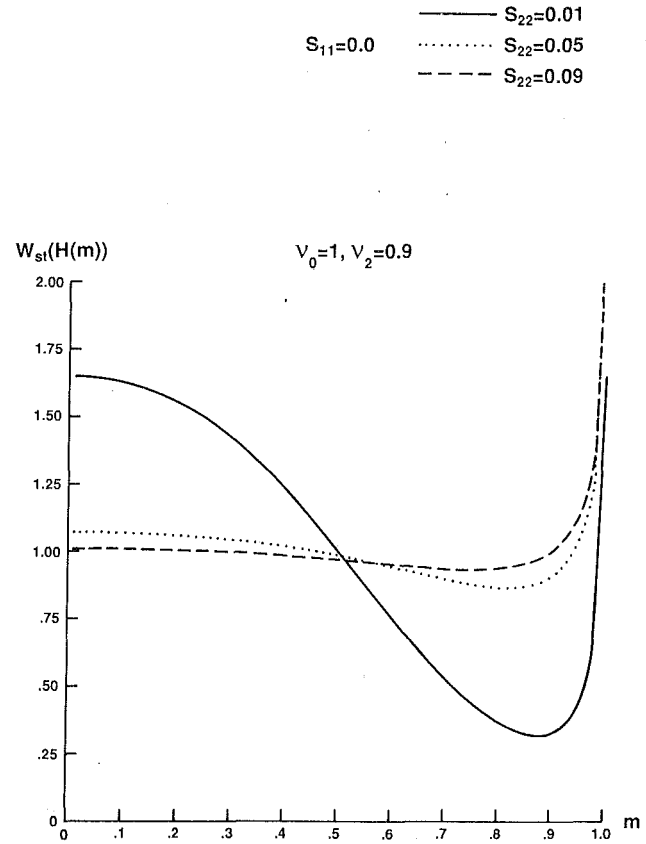


Fig. 1(*a*) The stationary density versus *m* for a parametrically perturbed system

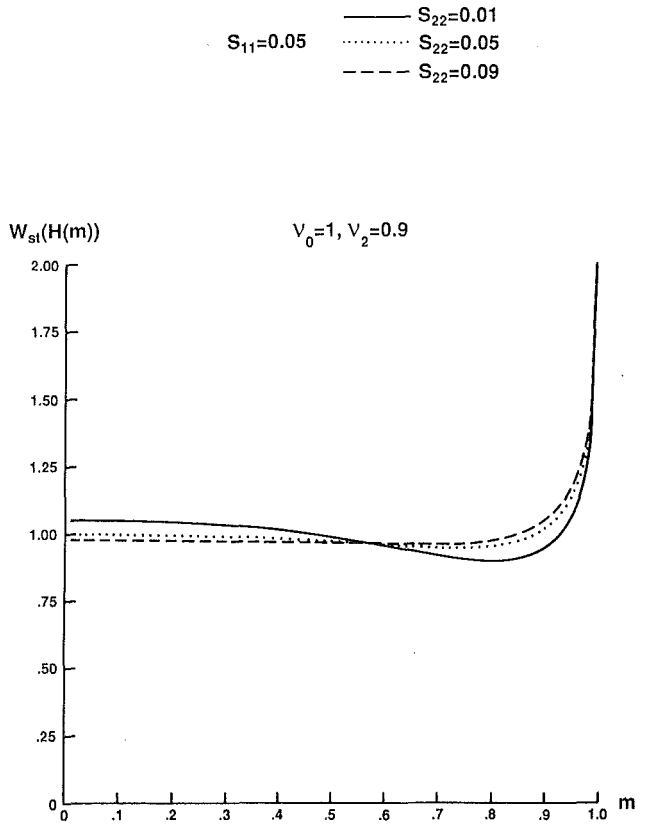


Fig. 1(*b*) The stationary density versus *m* for a system with both parametric and external excitations

$$-2[5(S_{22}/\nu_o)(2-m)(1-m) + S_{11}\nu_0^2(2-m)(1-m)]F(m)\}$$

$$\Lambda(H) = \int_{u_o^-(H)}^{u_o^+(H)} \frac{du}{\sqrt{Q_{1,2}(u)}} = \frac{(2-m)^{1/2}}{\sqrt{\nu_o}} F(m), \quad T(H) = 2\Lambda(H)$$

$$H = \frac{\nu_o^2(m-1)}{(2-m)^2}, \text{ and } 2C(H) = \frac{d}{dH}[\sigma_{HH}^2].$$

In the above equations, $F(m)$ and $E(m)$ are complete elliptic integrals of the first and second kinds, respectively. By solving the corresponding Fokker-Planck equation, the stationary probability density is obtained as

$$w_{st}(H) = \text{const. } \exp\{\psi\} \tag{7}$$

where

$$\psi = 2\int \frac{B(H)}{\sigma_{HH}^2(H)} dH + \ln(\Lambda(H)).$$

The stationary probability density functions are shown in Fig. 1 for various values of excitations. In Fig. 1(*a*), the parametric excitation is varied while keeping the external excitation fixed and vice versa in Fig. 1(*b*).

Following the arguments of Stratonovich (1963), the probability that the displacement is u at time t given the value of H, i.e., p(u, t |H), is proportional to the time a system spent at u knowing that the energy level is H. Furthermore, the time $u(t)$ spends at the point u is inversely proportional to the velocity and thus

$$w(u|H) = \frac{1}{T(H)} \frac{1}{\sqrt{Q_{1,2}(u)}} \text{ and}$$

$$w(u, H) = w(H)\ W(u|H)$$

and the density in u and v can be written as

$$w(u, v) = \text{const. } \exp\left[2\int_{H_o}^{H(u,v)} \{B(y)/\sigma_{HH}^2(y)\}dy\right]. \tag{8}$$

For the associated deterministic problem, i.e., $\eta_1(t) = \eta_2(t) = 0$, $\partial H/\partial t = 0$ gives constant energy loops such as limit cycles and separatrix, and the corresponding parameter values μ_o and μ_2 which satisfy $B(H) = 0$, can be obtained for values of $H \in (0, \nu_0^2/4)$ and $H \in (-\nu_0^2/4, 0)$, respectively, for case (i) and case (ii). As before, the extrema of $W(H)$ given by $\partial\psi/\partial H = 0$ denote, so to speak, the continuation of the deterministic constant energy levels or the "limit cycles" and are given by

$$2\Lambda(H)B(H) + \sigma_{HH}^2(H)\frac{d\Lambda}{dH}(H) = 0. \tag{9}$$

The stability of such limit cycles is determined by the sign of $(\partial^2\psi/\partial^2 H)_{H=H_o}$. These results are shown in Fig. 2, where the left-hand side of equation (9) is ploted against m and the most probable values are given by point m_c. In addition to these, the "fixed points" of the stochastic system are obtained by solving $\partial\psi/\partial u = 0$ and $\partial\psi/\partial v = 0$, and their stability is determined by the matrix of the second derivatives of ψ.

Exit Time Problems

In order to examine the stochastic stability of the equilibrium points from the Itô equation of the Hamiltonian H, it is important to determine the domain of attraction of the deterministic system. Knowing the domain of attraction we can say that the stochastic system losses stability (w.p.1) when the trajectories cross the boundary of this domain. In the sequel, the domain of attraction of the deterministic system is

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

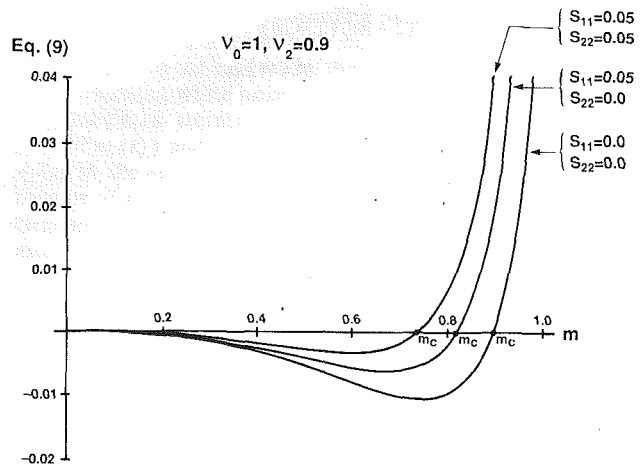


Fig. 2(a) The comparison of critical points of the stochastic system with that of the deterministic system

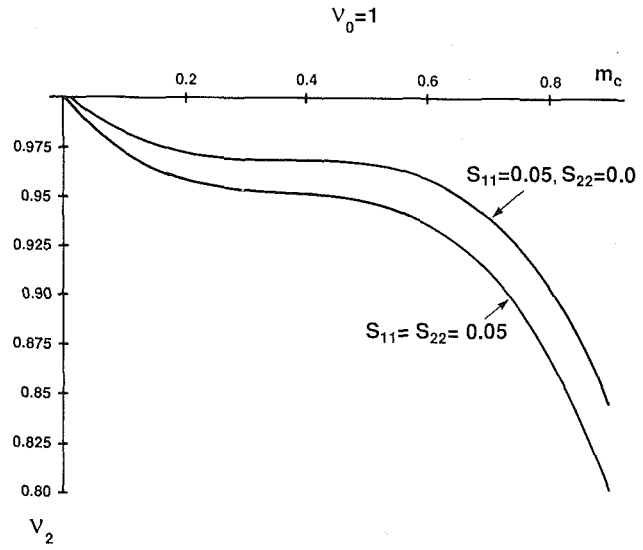


Fig. 2(b) The critical points versus ν_2 for the stochastic case

calculated for some regions in the parameter space for the case $P_2(\nu_0, u) = -(\nu_0^2/2)u + u^4/4$.

It is obvious when $\nu_o < 0$ and $\nu_2 < 0$, the trivial equilibrium point is asymptotically stable and when $\nu_0 < 0$ and $\nu_2 = 0$, the deterministic system undergoes a Hopf bifurcation and a stable limit cycle exists for $\nu_0 < 0$ and $\nu_2 > 0$. Similarly, when $\nu_2 < 0$ and $\nu_0 = 0$, the deterministic system undergoes a simple bifurcation and two stable equilibrium points exist for $\nu_2 < 0$ and $\nu_o > 0$. Furthermore, applying Bendixson's criterion, i.e., if the divergence of the vector field in R^2 has a fixed sign (zero excluded) in a simply connected region D in R^2 the system has no closed orbits lying entirely in D, we have $(\nu_2 - u^2) < 0$ for $\nu_2 < 0$ implying no closed orbits encircling all three fixed points. The associated stochastic problems for Hopf and simple bifurcations were discussed by the author (1988a, 1989). In this paper, attention is focussed on the region $\nu_o > 0$ and $\nu_2 > 0$.

Every periodic orbit for the deterministic system must encircle one or all three equilibrium points crossing the u-axis at $(b, 0)$. *If $b \epsilon (0, \sqrt{\nu_o})$, then the periodic orbit is a limit cycle encircling* the equilibrium point $(\sqrt{\nu_o}, 0)$. Due to symmetry, i.e., $P_2(u) = -P_2(-u)$, there exists another periodic orbit encircling the equilibrium point $(-\sqrt{\nu_o}, 0)$. Let the periodic orbit for the perturbed deterministic system be denoted as $\Gamma_\epsilon\ (b_1, \nu_0, \nu_2)$. Along the solutions of equation (4), with $\eta_1 = \eta_2 = 0$, we have

$$\dot{H} = \epsilon v^2 G(\nu_2, u), \tag{10}$$

and since $\Gamma_\epsilon\ (b, \nu_0, \nu_2)$ is a closed path, we also have

$$\int_{\Gamma_\epsilon} \dot{H} dt = 0, \quad \text{i.e.,}$$

$$\tilde{F}(b, \epsilon, \nu_0, \nu_2) = \int_{\Gamma_\epsilon} v^2 G(\nu_2, u) dt = 0. \tag{11}$$

The function $\tilde{F}(b, 0, \nu_0, \nu_2) \equiv \tilde{B}(b)$ may be written explicitly as

$$\tilde{B}(b) = \nu_2 \tilde{J}_o(b, \nu_o) - \tilde{J}_1(b, \nu_0) \tag{12}$$

where $\tilde{J}_o = \int_{\Gamma_o} v^2 dt$ and $\tilde{J}_1 = \int_{\Gamma_o} v^2 u^2 dt$.

Thus, the solution of $\tilde{F}(b, 0, \nu_o, \nu_2) \equiv \tilde{B}(b) = 0$ is given by

$$\nu_2 = \tilde{J}_1(b, \nu_o)/\tilde{J}_o(b, \nu_o). \tag{13}$$

Differentiating equation (12) yields

$$\frac{\partial \tilde{F}}{\partial \nu_2}(b, 0, \nu_o, \nu_2) = \tilde{J}_o(b, \nu_o) \neq 0$$

which implies, by the implicit function theorem (IFT), that there exists a unique continuously differentiable function $\nu^*(b, \epsilon, \nu_o)$ such that $\tilde{F}(b, \epsilon, \nu_0, \nu^*(b, \epsilon, \nu_o)) = 0$ for sufficiently small ϵ and

$$\nu^*(b, 0, \nu_o) = \tilde{J}_1(b, \nu_o)/\tilde{J}_o(b, \nu_o). \tag{14}$$

Having shown the existence of a limit cycle by IFT, we next proceed to list various periodic orbits present in the region $\nu_2 > 0$ and $\nu_o > 0$. However, *it is convenient to employ in place* of b another parameter H, which corresponds to the energy level $H = -(b^2/2)(\nu_o - b^2/2)$. This change of parameters is justified since $\partial H/\partial b = -b(\nu_o - b^2)$ is zero only at $b = \pm\sqrt{\nu_o}$ and 0 which are the fixed points. Thus, for $H \epsilon [-\nu_o^2/4, 0]$, making $B(H) \equiv 0$ yields

$$\nu_2(H, \nu_o) = J_1(H)/J_o(H) = R(H(m)) \cdot \nu_0 \tag{15}$$

where

$$J_o(H) = \int_{u_o^-(H)}^{u^+(H)} \left(2H + \nu_o u - \frac{U^4}{2}\right)^{1/2} du$$

$$= (2/3)[\nu_o/(2-m)]^{3/2}[(2-m)E(m) - 2(1-m)F(m)]$$

$$J_1(H) = \int_{u_o^-(H)}^{u^+(H)} \left(2H + \nu_o u - \frac{u^4}{2}\right)^{1/2} u^2 du$$

$$= (4/15)[\nu_o/(2-m)]^{5/2}[2(m^2 - m + 1)E(m) + (2-m)(m-1)F(m)]$$

and

$$m = 2(\nu_0^2 + 4H)^{1/2}/[\nu_0 + (\nu_0^2 + 4H)^{1/2}],\ m \epsilon [0, 1].$$

The following limits

$$\lim_{m \to 0} \nu_2(H(m), \nu_o) = \nu_o \text{ and } \lim_{m \to 1} \nu_2(H(m), \nu_o) = (4/5)\nu_o \tag{16}$$

agree with the calculations of Hopf bifurcation (local analysis) and saddle-loop (using Melnikov integrals), respectively, in the region $\nu_0 > 0$ and $\nu_2 > 0$ (see, for example, Guckenheimer and Holmes 1983). Furthermore, it can be shown as in Carr (1981) that the limit cycle encircling each of the nontrivial fixed points is unique for $H \epsilon [-\nu_o^2/4, 0]$.

It should be further noted that as the noise terms tend to zero, the extrema given by equation (9) tend to the steady-state solution of the deterministic system given by equation (15). This further emphasizes the fact that the most probable values are, so to speak, the continuation of the deterministic steady states. The critical values of both equations (9) and (15) are

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

given in Fig. 2. However, in this paper the domain of attraction is obtained using the critical points of the deterministic equation (5) which are shown in Fig. 3. It can be seen that the noise terms lower the critical values.

Now that we have established the domain of attraction, we formulate the exit time problem associated with the nontrivial fixed points in the region $\nu_2>0$, $\nu_o>0$ and determine the probabilistic information concerning the time T when the stochastic response process first passes out of a local domain of attraction of the fixed points. Since there are two simply connected domains for $H\epsilon[-\nu_o^2/4, 0]$, we shall concentrate on the region $\nu_2\epsilon[\nu_o, (4/5)\nu_o]$. Furthermore, for a specific value of ν_2 and ν_o, there is a unique value of H or, equivalently, $m\epsilon[0, 1]$ given by $m=R^{-1}(\nu_2/\nu_o)$, which defines the boundary Γ of the domain (see Fig. 3).

$$dm=\mu_m(m)dt+\sigma_{mm}(m)dw \tag{17}$$

where $\mu_m(m)=\left(\frac{1}{\Gamma(m)}\right)\left[B(m)+\frac{1}{2}\frac{d}{dm}[\sigma^2(m)]\right]$,

$$\Gamma(m)=\frac{m{\nu_0}^{3/2}F(m)}{(2-m)^{5/2}}$$

$$\sigma_{mm}^2(m)=\left(\frac{1}{\Gamma(m)}\right)[\sigma^2(m)],$$

$$\sigma^2(m)=\sigma_{HH}^2\frac{(2-m)^3}{m{\nu_o}^2},\ 0\leqq m\leqq 1.$$

Suppose that at time $t=0$, the state of the system corresponds to some point defined by m_o within D. When the random disturbance is applied, we are interested in the time T it takes for a trajectory at m_o to reach the boundary Γ of D for the first time, i.e.,

$$T=\min\{t:\ m(t)\epsilon\Gamma\,|m(0)=m_o\},\ m_o\epsilon D. \tag{18}$$

Now, defining the probability that a trajectory has not reached the boundary Γ during time interval τ as

$$P(\tau, m_o)=P_r\{\tau<T(m_o)\}, \tag{19}$$

the Kolmogorov's (backward) equation is written as

$$\frac{\partial P}{\partial\tau}(\tau, m_o)=\mu_m(m_o)\frac{\partial P}{\partial m_o}+\frac{1}{2}\sigma_{mm}^2(m_o)\frac{\partial^2 P}{\partial m_o}=L[P(\tau, m_o)] \tag{20}$$

with the initial and boundary conditions

$$P(0, m_o)=1\ \ m_o\epsilon D,\ P(\tau, m_c)=0\ \ m_c\epsilon\Gamma$$

The distribution function of the first passage time is $P_r[\tau=T]=1-P(\tau, m_o)$ and the corresponding Pontriagin equation for the nth moment is given by

$$L[M_n(m_o)]=-nM_{n-1}(m_o)\ \text{ and }\ M_n(m_c)=0. \tag{21}$$

Finally, the mean first passage time can be written as the solution of

$$L[M_1(m_o)]=\frac{1}{2}\sigma_{mm}^2(m_o)\frac{\partial^2 M_1}{\partial {m_o}^2}(m_o)+\mu_m(m_o)\frac{\partial M_1}{\partial m_o}(m_o)=-1 \tag{22}$$

with $M_1(m_c)=0$.

In addition to the boundary condition $M_1(m_c)=0$, a boundedness condition at $m_o=0$, $M_1(0)<\infty$, is usually imposed on the solution to uniquely determine $M_1(m_o)$. This condition, by itself, implies that the left boundary $m_o=0$ is not an absorbing boundary. This condition may be violated if the noise term $\sigma_{mm}^2(m_o)$ vanishes at $m_o=0$, and cannot then be used to obtain the solution. It is, therefore, important to

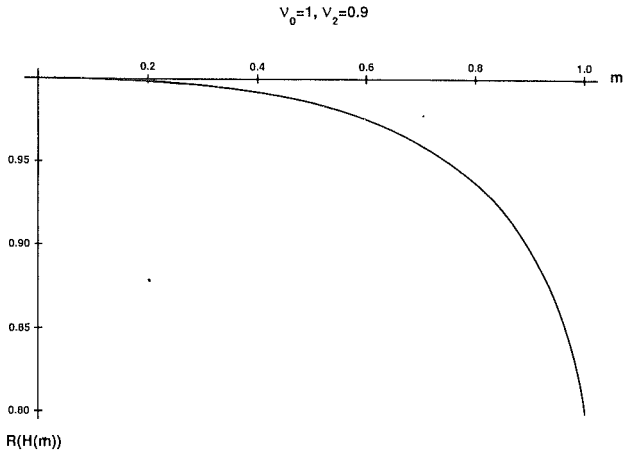


Fig. 3 The right boundary of the domain of attraction versus ν_2

understand the behavior of the diffusion process $m(t)$ near the boundary $m_o=0$ according to various Feller classifications (see, for example, Feller, 1954; Itô and McKean, 1964; Karlin and Taylor, 1981). To this end, consider an interval $(0, m_c]$ and let $\Delta\epsilon(0, m_c)$ be an interior point such that $m(t)$ is a regular diffusion process in $[\Delta, m_c]$. Putting T_x as the hitting time of x, we can define $\bar{M}_1(m_o)$ as mean time to reach either Δ or m_c, i.e.,

$$\bar{M}_1(m_o)=E\{T_\Delta\wedge T_{m_c}\,|m(0)=m_o\}\ \text{ for }\ \Delta<m_o<m_c$$

and can be evaluated as

$$\bar{M}_1(m_o)=2[1-\alpha(m_o, m_c)/\alpha(\Delta, m_c)]\cdot\int_{m_o}^{m_c}\Gamma(\eta)\alpha(\eta, m_c)\exp[\Phi(\eta)]d\eta+2\alpha(m_o, m_c)/\alpha(\Delta, m_c)\int_\Delta^{m_o}\Gamma(\eta)\alpha(\Delta, \eta)\exp[\Phi(\eta)]d\eta \tag{23}$$

where

$$\Phi(\eta)=\int_o^\eta\frac{2B(x)}{\sigma^2(x)}dx,\ \alpha(z_1, z_2)=\int_{z_1}^{z_2}\exp[-\Phi(x)]\frac{dx}{\sigma^2(x)}.$$

In order to classify the boundary behavior, consider the following quantities γ_o and β_o, which can be respectively defined as roughly the measure of time to reach the left boundary 0 starting from an interior point $m_o\epsilon(0, m_c)$, and the measure of time to reach an interior point m_o starting from the left boundary 0. They are defined as

$$\gamma_o=\int_o^z\frac{1}{\sigma^2(x)}\left\{\int_x^z\Gamma(\eta)\exp[\Phi(\eta)]d\eta\right\}\exp[-\Phi(x)]dx \tag{24}$$

$$\beta_o=\int_o^z\Gamma(\eta)\left\{\int_\eta^z\exp[-\Phi(x)]\frac{dx}{\sigma^2(x)}\right\}\exp[\Phi(\eta)]d\eta \tag{25}$$

where z is an interior point. The Feller classification of the left boundary, 0, in terms of γ_o and β_o is as follows:

(1) The left boundary 0 is *regular* if $\gamma_o<\infty$ and $\beta_o<\infty$. The process can both enter and leave from the boundary 0. In other words, the process starting from an interior point can reach the boundary with some positive probability in finite time. Similarly, the process starting from the boundry can reach an interior point with some positive probability in finite time.

(2) The left boundary 0 is an *exit* if $\gamma_o<\infty$ and $\beta_o=\infty$. The process starting from an interior point can reach the boundary with some positive probability, but starting at 0, it is impossi-

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

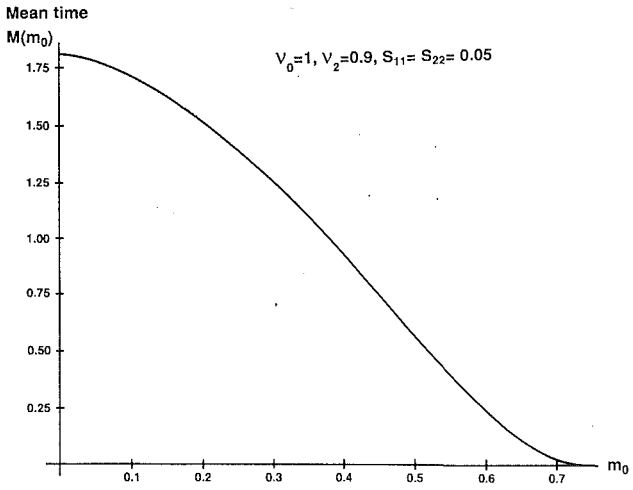


Fig. 4 The mean exit time versus m_o

ble to reach any interior point $m_o \epsilon (0, m_c)$. Furthermore, the exit boundary is either a trap or an absorbing point.

(3) The left boundary 0 is an *entrance* if $\gamma_o = \infty$ and $\beta_o < \infty$. An entrance boundary cannot be reached from any interior point; i.e., the probability is zero that the process starting at an interior point can reach an entrance boundary. Furthermore, the process starting from an entrance boundary moves at once to the interior never to return to the boundary.

(4) The left boundary 0 is *natural* if $\gamma_o = \infty$ and $\beta_o = \infty$. The process starting from an interior point cannot reach the boundary in finite time and the process cannot reach any interior point starting from the natural boundary.

Using this classification, Kozin and Sunahara (1987) established some stability properties of the singular point 0. Along similar lines, we establish an instability condition for the singular point 0.

Theorem: Let the singular point 0 be the left boundary of the interval $(0, m_c]$. The singular point is unstable in probability; i.e., the event $\{\sup |m(t{:}m_o, 0)| < \epsilon\}$ has probability zero for all $m_o > 0$, if the following conditions are satisfied

(i) 0 is an entrance boundary, i.e., $\gamma_o = \infty$ and $\beta_o < \infty$

(ii) m_c is a regular or an exit boundary, i.e., $\gamma_{m_c} < \infty$

where

$$\gamma_{m_c} = \int_z^{m_c} \frac{1}{\sigma^2(x)} \left\{ \int_z^x \Gamma(\eta) \exp[\Phi(\eta)] d\eta \right\} \exp[-\Phi(x)] dx. \quad (26)$$

Moreover, the mean time to reach m_c is given by

$$M_1(m_o) = 2 \int_{m_o}^{m_c} \Gamma(\eta) \left\{ \int_\eta^{m_c} \exp \right.$$

$$\left. [-\Phi(x)] \frac{dx}{\sigma^2(x)} \right\} \exp[\Phi(\eta)] d\eta \quad (27)$$

and $M_1(0) < \infty$.

Proof: The proof follows the boundary behavior according to the above classification. From the definition of β_o it is obvious that $\beta_o < \infty$ implies

$$\lim_{\Delta \downarrow 0} \int_\Delta^z \Gamma(\eta) \exp[\Phi(\eta)] d\eta < \infty, \quad \Delta < z < m_c$$

and hence for an entrance boundary the following relationship

$$\gamma_o + \beta_o = \lim_{\Delta \downarrow 0} \left\{ \alpha(\Delta, z) \int_\Delta^z \Gamma(\eta) \exp[\Phi(\eta)] d\eta \right\}$$

yields $\lim_{\Delta \downarrow 0} \alpha(\Delta, z) = \infty$. Thus, for some m_o, $0 < m_o < z < m_c$, the $\Pr\{T_z < T_{o+} \mid m(0) = m_o\}$ is given by

$$\lim_{\Delta \downarrow 0} [1 - \alpha(m_o, z)/\alpha(\Delta, z)] \to 1.$$

Moreover, since $\beta_o < \infty$, we conclude that the process starting at any $m_o \epsilon [o, z]$ reaches z in finite time with probability one. If the right boundary is regular, the z is replaced by m_c. The right boundary is attractive and attainable when m_c is an exit point. Thus, for both these cases

$$\Pr\{\sup_{t>0} |m(t{:}m_o, 0)| < \epsilon\} = 0, \text{ for all } m_o > 0.$$

Now, using the fact $\lim_{\Delta \downarrow 0} \alpha(\Delta, m_c) = \infty$, the exit time can be obtained from equation (23) as well as (27).

For the problem under consideration, numerical calculations indicate that $\gamma_o = \infty$, $\beta_o < \infty$ and $\gamma_{m_c} < \infty$ and thus the singular point 0 is unstable in probability. The mean exit time to reach m_c can be calculated using equation (27). These results are shown in Fig. 4.

Conclusions

In this paper, a special case of co-dimension two stochastic bifurcation associated with non-semi-simple double-zero eigenvalues is examined. The normal form associated with this case is two dimensional and corresponds to the stochastically perturbed van der Pol-Duffing equation. The stochastic averaging method appropriate for the Itô equations is applied to obtain a one-dimensional Itô equation for the Hamiltonian. The probability density and its extrema are obtained. It is found that the extrema of the density function correspond to the least probable value from which the process leaves rather quickly. The critical extrema values m_c of the stochastic system approach the values of m at which the deterministic unstable limit cycle exists as the noise terms tend to zero. Furthermore, the effect of additive and multiplicative noise terms on the values of m_c are demonstrated and it is shown that the noise terms lower these critical values. In order to examine the stochastic stability of the equilibrium points from the Itô equation for the Hamiltonian, the domain of attraction of the deterministic system is obtained. The stability in probability and the mean exit time are calculated using the domain of attraction. The stability in probability is obtained by examining the boundary behavior of the one-dimensional diffusion process. An instability theorem is stated and proven. It is found that the nontrivial fixed points are unstable in probability. The mean exit time to reach the boundaries of the domain of attraction is determined for various noise levels.

Acknowledgment

This research is supported by a grant from the Air Force Office of Scientific Research under Grant No. AFOSR-88-0233. Dr. Spencer Wu is the Program Director.

References

Arnold, L., Horsthemke, W., and Lefever, R., 1973, "White and Colored External Noise and Transition Phenomena in Nonlinear Systems," *Z. Physik*, Vol. B29, pp. 367–373.

Arnold, L., Horsthemke, W., and Stucki, J. W., 1979, "The Influence of External Real and White Noise on the Lotka-Voltera Model," *Biom. J.*, Vol. 21, No. 50, pp. 451–471.

Baras, F., Mansour, Malek M., and Van den Broenck, C., 1982, "Asymptotic Properties of Coupled Nonlinear Langevin Equations in the Limit of Weak Noise, II: Transition to a Limit Cycle," *J. Stat. Phys.*, Vol. 28, No. 3, pp. 577–587.

Carr, J., 1981, *Applications of Center Manifold Theory*, Springer-Verlag, New York.

Feller, W., 1954, "Diffusion Process in One Dimension," *Trans. of Amer. Math. Soc.*, Vol. 97, pp. 1–31.

Graham, R., 1982, "Hopf Bifurcations with Fluctuating Control Parameter," *Phys. Rev.*, Vol. A25, pp. 3234–3258.

Graham, R., 1987, "Macroscopic Potentials, Bifurcations and Noise in

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Dissipative Systems," *Fluctuations and Stochastic Phenomena in Condensed Matter* (Lecture Notes in Physics, 268), L. Garrido, ed., Springer-Verlag, New York.

Guckenheimer, J., and Holmes, P., 1983, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York.

Horsthemke, W., and Lefever, R., 1977, "Phase Transition Induced by External Noise," *Phys. Lett.*, Vol. A64, pp. 19–23.

Ito, K., and McKean, H. P., 1964, *Diffusion Processes and Their Sample Paths*, Academic Press, New York.

Karlin, S., and Taylor, H. M., 1981, *A Second Course in Stochastic Processes*, Academic Press, New York.

Khasminskii, R. Z., 1967, "Necessary and Sufficient Conditions for the Asymptotic Stability of Linear Stochastic Systems," *Theo. Prob. Appl.*, Vol. 12, pp. 167–172.

Khasminskii, R. Z., 1968, "On the Principles of Averaging for Itô Stochastic Differential Equations," *Kybernetica*, Vol. 4, pp. 260–279.

Kozin, F., and Prodromou, S., 1971, "Necessary and Sufficient Conditions for Almost Sure Sample Stability of Linear Ito Equations," *SIAM J. Appl. Math.*, Vol. 21, pp. 413–424.

Kozin, F., and Sunahara, Y., 1987, "An Application of the Averaging Method to Noise Stabilization of Nonlinear Systems," *Proc. 20th Midwestern Conference*, West Lafayette, Ind, Vol. 14(a), pp. 291–298.

Lefever, R., and Turner, J. W., 1984, *Non-Equilibrium Dynamics in Chemical Systems*, C. Vidal and A. Pacault, eds., Springer-Verlag, Berlin.

Nishioka, K., 1976, "On the Stability of Two Dimensional Linear Stochastic Systems," *Kodai Math. Sem. Rep.*, Vol. 27, pp. 211–230.

Sri Namachchivaya, N., 1988, "Hopf Bifurcation in the Presence of both Parametric and External Stochastic Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 923–930.

Sri Namachchivaya, N., 1990, "Stochastic Bifurcations," *J. Appl. Math. and Computations*, Vol. 38, pp. 101–159.

Sri Namachchivaya, N., and Lin, Y. K., 1988, "Applications of Stochastic Averaging for Systems with High Damping," *Prob. Eng. Mech.*, Vol. 3, No. 3, pp. 159–167, see also *Stochastic Structural Mechanics*, Lecture Notes in Engineering, Y. K. Lin and G. I. Schueller, eds., 1987, Springer-Verlag, New York, pp. 282–310.

Stratonovich, R. L., 1963,, *Topics in the Theory of Random Noise*, Vols. 1 and 2, Gordon and Breach, New York.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Maximum Entropy Approach for Stationary Response of Nonlinear Stochastic Oscillators

R. J. Chang
Associate Professor,
Department of Mechanical Engineering,
National Cheng Kung University,
Tainan, Taiwan 70101
Republic of China

A new approach based on the maximum entropy method is developed for deriving the stationary probability density function of a stable nonlinear stochastic system. The technique is implemented by employing the density function with undetermined parameters from the entropy method and solving a set of algebraic moment equations from a nonlinear stochastic system for the unknown parameters. For a wide class of stochastic systems with given density functions, an explicit density function of the stochastic system perturbed by a nonlinear function of states and noises can be obtained. Three nonlinear oscillators are selected for illustrating the present scheme and the validity of the derived density functions is further supported by some exact solutions and Monte Carlo simulations.

1 Introduction

The investigation of the dynamic response of a nonlinear stochastic oscillator has been attracted numerous researchers in the areas of engineering mechanics and structural engineering. Several techniques have been developed for obtaining the stationary and nonstationary statistical response of a nonlinear oscillator under stochastic parametric and/or external excitations (Caughey, 1986; Cai and Lin, 1988; Crandall, 1985; Ibrahim and Heo, 1987; Liu and Huw, 1988; Robert and Spanos, 1986; Young and Chang, 1987). For a general nonlinear stochastic oscillator, the associated Fokker-Planck-Kolmogorov (FPK) equation with appropriate boundary and initial conditions which governing the propagation of probability density functions of states always can be formulated. However, even for the stationary response, the exact density function can be obtained only for some classes of stochastic nonlinear oscillators and is almost impossible to be derived for a general nonlinear stochastic system including parametric noise-excited terms. Therefore, only partial description of the response statistics such as the second moment response usually can be obtained in developing some approximate techniques (Ibrahim and Heo, 1987; Liu and Huw, 1988).

In recent years, much research effort has been devoted to the solution of the stationary probability density function of a nonlinear stochastic system (Caughey, 1986; Cai and Lin, 1988; Dimentberg, 1985). Although the exact density functions of some classes of practical nonlinear oscillators have been derived, there are a wide class of nonlinear oscillators which can be described as the perturbation of a stochastic oscillator with solvable exact density function. For example, the stationary probability density function of a linear oscillator under both external and weakly parametric noise excitations can be considered as a perturbed jointly Gaussian density. Therefore, it is essential to develop some techniques for extending existing exact density function of a stochastic oscillator to include the effect of some perturbed functions of states and noises in the area of nonlinear random mechanics.

In this paper, a method based on the maximum entropy approach is developed for deriving the stationary probability density function of a nonlinear stochastic system. By employing the entropy method, the existing exact density function of a nonlinear oscillator can be easily extended to include the effect of a perturbed nonlinear function of states and noises. For a wide class of nonlinear oscillators, an explicit approximate density function can be further derived by employing the present technique.

2 Problem Formulation

Consider an n-dimensional Ito's nonlinear stochastic system which is derived from replacing the physical white noise by an ideal one through incorporating Wong-Zaikai correction term (Ibrahim, 1985) and is described as

$$d\mathbf{x} = \mathbf{f}_d(\mathbf{x})dt + G_d(\mathbf{x})d\mathbf{w}_d + \epsilon\mathbf{f}_p(\mathbf{x})dt + \epsilon^{\frac{1}{2}}G_p(\mathbf{x})d\mathbf{w}_p \qquad (1)$$

where $\mathbf{x}$ is an $n\times 1$ vector of states process, $\mathbf{f}_d$ is a dominant $n\times 1$ vector nonlinear function of states, $G_d \cdot d\mathbf{w}_d$ is a dominant $n\times 1$ noise excited part, ϵ is a constant, $\mathbf{f}_p$ is a perturbed $n\times 1$ vector nonlinear function of states, and $G_p \cdot d\mathbf{w}_p$ is a perturbed $n\times 1$ vector noise excited part. The random excitations are assumed to be mutually independent Wiener processes with independent increments and with intensities

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

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 10, 1989; final revision, May 4, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$E[d\mathbf{w}_d(t)\,d\mathbf{w}_d^T(t)] = Q_d dt$$
$$E[d\mathbf{w}_p(t)\,d\mathbf{w}_p^T(t)] = Q_p dt. \tag{2}$$

The problem is to derive the stationary probability density function of states of a nonlinear stochastic system described by equations (1) and (2). If the value of ϵ is small and the exact probability density function of the nonlinear system with dominant part in equation (1) by setting $\epsilon = 0$ is given, we want to derive an explicit approximate density function for the nonlinear system given by equation (1). The proposed technique for deriving the stationary probability density function will be developed through the maximum entropy formulation.

3 Formulation of Maximum Entropy Method

The maximum entropy methods have been successfully applied to the spectral estimations (Burg, 1975) and first-excursion failure problems (Lin, 1970); however, the techniques have not been used for obtaining the output density function of a nonlinear stochastic system. The formulation of maximum entropy method for deriving the stationary probability density function of equation (1) will be proceeded by the following procedure. First, the maximum entropy principle will be introduced. In the information theory, the Shannon's entropy is defined as (Tribus, 1969)

$$H = -\int_R p^t(\mathbf{x}) \ln(\mathbf{x})\, d\mathbf{x} \tag{3}$$

where $p^t(\mathbf{x})$ is the joint probability density function of states $\mathbf{x}$ and R is an n-dimensional real domain for $\mathbf{x}$. *A priori* information about $\mathbf{x}$ is assumed given by

$$\int_R x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n} p^t(\mathbf{x})\, d\mathbf{x} = a_{l_1 l_2 \ldots l_n},$$
$$l_1, l_2, \ldots l_n = 0,1,2, \ldots n. \tag{4}$$

From Jaynes' principle (Tribus, 1969) stated as "The minimally prejudiced probability distribution is that which maximizes the entropy subjected to constraints supplied by the given information," one can derive $p^t(\mathbf{x})$ such that the entropy given by equation (3) will be maximized subjected to the constraint equations (4). Introducing Lagrange multipliers $\lambda_{l_1 l_2 \ldots l_n}$ to associate equations (3) and (4) and from the variational scheme, one derives

$$p^t(\mathbf{x}) = \exp\left[\sum_{l_1=0}^{m_1}\sum_{l_2=0}^{m_2}\cdots\sum_{l_n=0}^{m_n} \lambda_{l_1 l_2 \ldots l_n} x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n} - 1\right]. \tag{5}$$

By absorbing the constant term of equation (5) in N, then (5) can be expressed as

$$p^t(\mathbf{x}) = N \exp\left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_n} \lambda_{l_1 \ldots l_n} x_1^{l_1} \cdots x_n^{l_n}\right] \tag{6}$$
$$\text{except all } l_i = 0$$

where N is a normalization constant given by

$$N = \frac{1}{\int_R \exp\left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_n} \lambda_{l_1 \ldots l_n} x_1^{l_1} \cdots x_n^{l_n}\right] d\mathbf{x}}. \tag{7}$$
$$\text{Except all } l_i = 0.$$

The probability density function of $\mathbf{x}$ can be constructed as equation (6) if the undetermined parameters $\lambda_{l_1 l_2 \ldots l_n}$ can be obtained through equation (1). Now, from equations (1) and (2), the associated stationary FPK equation can be formulated as

$$-\sum_{i=1}^{n} \frac{\partial}{\partial x_i}[p^t(\mathbf{x})(f_d^i + \epsilon f_p^i)] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}[p^t(\mathbf{x})(G_d(\mathbf{x}) Q_d G_d^T(\mathbf{x}))_{ij}] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}[p^t(\mathbf{x})\epsilon(G_p(\mathbf{x}) Q_p G_p^T(\mathbf{x})_{ij}] = 0. \tag{8}$$

By multiplying equation (8) by differentiable functions $\phi(\mathbf{x})$ and integrating the resulting equation over the state space, the moment relations can be derived to give

$$\sum_{i=1}^{n} E\left[\frac{\partial \phi(\mathbf{x})}{\partial x_i}(f_d^i + \epsilon f_p^i)\right] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left[\frac{\partial^2 \phi(\mathbf{x})}{\partial x_i \partial x_j}(G_d(\mathbf{x}) Q_d G_d^T(\mathbf{x}) + \epsilon G_p(\mathbf{x})\, Q_p G_p^T(\mathbf{x}))_{ij}\right] = 0. \tag{9}$$

By employing equation (6) and selecting appropriate functions $\phi(\mathbf{x})$ in equation (9) for evaluating the expected values, the unknown $\lambda_{l_1 l_2 \ldots l_n}$ in equation (6) can be derived if $(m_1+1) \cdot (m_2+1) \cdot \ldots \cdot (m_n+1) - 1$ independent moment relations are formulated.

The formulation of maximum entropy method for solving the stationary probability density function of equation (1) can be preceded by the above procedure. However, the evaluation of expected value by employing the density function of equation (6) usually cannot be derived explicitly. Consequently, a numerical integration algorithm implemented as a weighted summation scheme needs to be employed for evaluating the expected value in equation (9) (Siddall and Diab, 1975). Hence, a set of nonlinear algebraic equations can be formulated from equation (9) for solving the undetermined Lagrange multipliers and the stationary probability density function will be given by equation (6).

If the value of ϵ in equation (1) is small, the present scheme will be further simplified for solving the approximate stationary density function. For the nonlinear stochastic system given by equation (1), one may assume that the stationary probability density function which is described as equation (6) can be further expressed as

$$p^t(\mathbf{x}) = N \exp\left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_n} (\lambda_{l_1 \ldots l_n}^d + \epsilon \lambda_{l_1 \ldots l_n}^p) x_1^{l_1} \cdots x_n^{l_n}\right].$$
$$\text{Except all } l_i = 0. \tag{10}$$

By expanding equation (10) in a Taylor's series and retaining the first order of ϵ in the expansion, one has

$$p^t(\mathbf{x}) = N\, p_d(\mathbf{x})\left[1 + \epsilon \sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_2} \lambda_{l_1 \ldots l_n}^p x_1^{l_1} \cdots x_n^{l_n}\right]$$
$$\text{except all } l_i = 0 \tag{11}$$

where

$$p_d(\mathbf{x}) = \exp\left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_n} \lambda_{l_1 \ldots l_n}^d x_1^{l_1} \cdots x_n^{l_n}\right]$$
$$\text{except all } l_i = 0. \tag{12}$$

If one defines

$$p_d^t(\mathbf{x}) = N_d p_d(\mathbf{x}) \tag{13}$$

and

$$p_p^t(\mathbf{x}) = N_p p_d(\mathbf{x}) \sum_{l_1=0}^{m_1} \cdots \sum_{l_n=0}^{m_n} \lambda_{l_1 \ldots l_n}^p x^{l_1} \cdots x^{l_n}. \tag{14}$$

Except all $l_i = 0$.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with

$$\int_R p_d(\mathbf{x})d\mathbf{x}=\frac{1}{N_d} \tag{15}$$

and

$$\int_R p_d(\mathbf{x})\sum_{l_1=0}^{m_1}\cdots\sum_{l_n=0}^{m_n}\lambda^p_{l_1\ldots l_n}x_1^{l_1}\cdots x_n^{l_n}d\mathbf{x}=\frac{1}{N_p}. \tag{16}$$

Except all $l_i=0$, then equation (11) becomes

$$p^t(\mathbf{x})=\frac{N}{N_d}p_d^t(\mathbf{x})+\epsilon\frac{N}{N_p}p_p^t(\mathbf{x}). \tag{17}$$

By employing equation (17) for $p^t(\mathbf{x})$ and utilizing equations (15) and (16), equation (7) becomes

$$N=\frac{1}{1/N_d+\epsilon/N_p}. \tag{18}$$

By utilizing equations (13), (14), and (17) in equation (9) for evaluating the expected value, one derives the following equations:

$$\epsilon^0:\sum_{i=1}^{n}E_d\left[\frac{\partial\phi(\mathbf{x})}{\partial x_i}f_d^i\right]+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}E_d\left[\frac{\partial^2\phi(\mathbf{x})}{\partial x_i\partial x_j}(G_d(\mathbf{x})Q_dG_d^T(\mathbf{x}))_{ij}\right]=0 \tag{19}$$

$$\epsilon^1:\sum_{l_1=0}^{m_1}\cdots\sum_{l_n=0}^{m_n}\lambda^p_{l_1\ldots l_n}\left\{\sum_{i=1}^{n}E_d\left[\frac{\partial\phi(\mathbf{x})}{\partial x_i}f_d^ix_1^{l_1}\cdots x_n^{l_n}\right]\right.$$

Except all $l_i=0$.

$$\left.+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}E_d\left[\frac{\partial^2\phi(\mathbf{x})}{\partial x_i\partial x_j}(G_d(\mathbf{x})Q_d\,G_d^T(\mathbf{x}))_{ij}x_1^{l_1}\cdots x_n^{l_n}\right]\right\}$$

$$+\sum_{i=1}^{n}E_d\left[\frac{\partial\phi(\mathbf{x})}{\partial x_i}f_p^i\right]+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}E_d\left[\frac{\partial^2\phi(\mathbf{x})}{\partial x_i\partial x_j}(G_p(\mathbf{x})Q_p\,G_p^T(\mathbf{x}))_{ij}\right]=0 \tag{20}$$

where $E_d[\cdot]$ is defined as

$$E_d[\cdot]=\int_R\cdot p_d^t(\mathbf{x})d\mathbf{x}. \tag{21}$$

By selecting appropriate $\phi(\mathbf{x})$, it is noted that equation (19) formulates the moment relations of the dominant part of equation (1), and equation (20) will formulate a set of linear algebraic equations with unknown $\lambda^p_{l_1l_2\ldots l_n}$ to account for the effect of the perturbed part of equation (1). Thus, the unknown $\lambda^p_{l_1l_2\ldots l_n}$ can be derived from equation (20) if $E_d[\cdot]$ can be evaluated by employing the joint probability density function of the dominant part of equation (1) obtained from equation (19) or from the associated *FPK* equation. As a result, the stationary probability density function of equation (1) can be expressed by employing equations (10) and (12) to yield

$$p^t(\mathbf{x})=N\,p_d(\mathbf{x})\exp\left(\epsilon\sum_{l_1=0}^{m_1}\cdots\sum_{l_n=0}^{m_n}\lambda^p_{l_1\ldots l_n}x_1^{l_1}\cdots x_n^{l_n}\right) \tag{22}$$

except all $l_i=0$

where N is given by

$$N=\frac{1}{\int_R p_d(\mathbf{x})\exp\left(\epsilon\sum_{l_1=0}^{m_1}\cdots\sum_{l_n=0}^{m_n}\lambda^p_{l_1}\ldots l_nx_1^{l_1}\ldots x_n^{l_n}\right)d\mathbf{x}}. \tag{23}$$

Except all $l_i=0$.

The formulation of the maximum entropy method has been presented in this section. The application of the present technique for deriving the stationary probability density function will be illustrated in the next section by selecting three examples.

4 Examples and Discussion

Example 1. Consider a Duffing oscillator given by

$$\ddot{x}+2\xi_0\omega_0\dot{x}+\mu_0x^3+\epsilon\omega_0^2x=w(t) \tag{24}$$

where

$$E[w(t)w(s)]=2q_0\delta(t-s). \tag{25}$$

For a stochastic nonlinear oscillator described as equation (24), the stochastic system can be expressed as equation (1) with $x_1=x$, $x_2=\dot{x}$ and with system parameters

$$\mathbf{f}_d=\begin{pmatrix}x_2\\-2\xi_0\omega_0x_2-\mu_0x_1^3\end{pmatrix}\quad\mathbf{f}_p=\begin{pmatrix}0\\-\omega_0^2x_1\end{pmatrix}$$

$$G_d=\begin{pmatrix}0\\1\end{pmatrix}\quad d\mathbf{w}_p=0$$

$$d\mathbf{w}_d=w(t)dt. \tag{26}$$

The stationary probability density function of x_1 and x_2 of the dominant part of equation (24), by setting $\epsilon=0$, can be derived from the associated *FPK* equation to yield two independent densities as

$$p_d^t(x_1)=\frac{1.682}{\Gamma(1/4)Q^{1/4}}\exp\left(-\frac{x_1^4}{2Q}\right) \tag{27}$$

where

$$Q=\frac{q_0}{\xi_0\omega_0\mu_0} \tag{28}$$

and

$$p_d^t(x_2)=\frac{1}{(2\pi E_d[x_2^2])^{1/2}}\exp\left(-\frac{x_2^2}{2E_d[x_2^2]}\right), \tag{29}$$

where

$$E_d[x_2^2]=\frac{q_0}{2\xi_0\omega_0}. \tag{30}$$

By substituting $\phi(\mathbf{x})=x_1^{r1}\,x_2^{r2}$ and equation (26) into equation (20) and employing the property that x_1 and x_2 are independent in evaluating $E_d[\cdot]$, one obtains

$$\sum_{l_1=0}^{m_1}\sum_{l_2=0}^{m_2}\lambda^p_{l_1l_20\ldots0}\{r_1E_d[x_1^{r1+l_1-1}]\,E_d\,[x_2^{r2+l_2+1}]$$

except all $l_i=0$.

$$-2\xi_0\omega_0r_2E_d\,[x_1^{r1+l_1}]E_d\,[x_2^{r2+l_2}]-\mu_0r_2E_d\,[x_1^{r1+l_1+3}]E_d\,[x_2^{r2+l_2-1}]$$
$$+q_0r_2(r_2-1)E_d[x_1^{r1+l_1}]E_d[x_2^{r2+l_2-2}]\}$$
$$-\omega_0^2r_2E_d\,[x_1^{r1+1}]E_d\,[x_2^{r2-1}]=0. \tag{31}$$

By choosing $m_1=m_2=2$ and ultilizing equations (27) to (30), the following equations can be obtained by selecting appropriate r_1 and r_2 as

$r_1=1$, $r_2=0$:

$$\lambda^p_{010\ldots0}=0 \tag{32}$$

$r_1=0$, $r_2=1$ and with equation (32):

$$\lambda^p_{100\ldots0}=0 \tag{33}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$r_1=2$, $r_2=0$ and with equations (32), (33):

$$\lambda^p_{110..0}=0 \tag{34}$$

$r_1=1$, $r_2=1$ and with equations (32) to (34):

$$(E_d[x_1^2]E_d[x_2^2]-\mu_0 E_d[x_1^6])\lambda^p_{200..0} + (E_d[x_2^4]-\mu_0 E_d[x_1^4]E_d[x_2^2])\lambda^p_{020..0}-\omega_0^2 E_d[x_1^2]=0 \tag{35}$$

$r_1=0$, $r_2=2$ and with equations (32) to (34):

$$(-2\xi_0\omega_0 E_d[x_1^2]E_d[x_2^2]+q_0 E_d[x_1^2])\lambda^p_{200..0} + (-2\xi_0\omega_0 E_d[x_2^4]+q_0 E_d[x_2^2])\lambda^p_{020..0}=0. \tag{36}$$

By substituting equation (30) into (36), one has

$$\lambda^p_{020..0}=0. \tag{37}$$

By employing equation (27) for evaluating the expected value $E_d[\cdot]$ in equation (35) and utilizing (30) and (37), one obtains

$$\lambda^p_{200..0}=-\frac{\omega_0^2}{Q\mu_0}. \tag{38}$$

Hence, the stationary probability density function of the perturbed stochastic system (24) can be expressed by substituting equations (27) to (30), (32) to (34), (37), and (38) into (22) to yield

$$p^t(\mathbf{x})=N\exp\left(-\frac{x_1^4}{2Q}-\epsilon\frac{\omega_0^2}{Q\mu_0}x_1^2\right)\exp\left(-\frac{x_2^2}{Q\mu_0}\right). \tag{39}$$

From equation (39), it is noted that the stationary probability density function of states derived from utilizing a given joint density function of the dominant part is exactly the same as the exact density function directly derived from equation (24) by employing *FPK* equation. Hence, by following the present approach, the stationary density function of a stochastic system can be directly extended to include the perturbed effect of a linear/nonlinear function of states. The next example selected for investigation is a second-order linear oscillator with a linear spring and noise-multiplicative term.

Example 2. A second-order stochastic parametric and external noise-excited oscillator is described as

$$\ddot{x}+2\xi_0\omega_0\dot{x}+\omega_0^2 x+\epsilon^{1/2}\mu(t)x=w(t) \tag{40}$$

with

$$E[\mu(t)\mu(s)]=2q_1\delta(t-s)$$

$$E[w(t)w(s)]=2q_0\delta(t-s). \tag{41}$$

The stochastic system can be formulated as equation (1) with system parameters

$$\mathbf{f}_d=\begin{pmatrix}x_2\\-2\xi_0\omega_0x_2-\omega_0^2x_1\end{pmatrix} \qquad \mathbf{f}_p=\begin{pmatrix}0\\0\end{pmatrix}$$

$$G_d=\begin{pmatrix}0\\1\end{pmatrix} \qquad G_p=\begin{pmatrix}0\\-x_1\end{pmatrix}$$

$$d\mathbf{w}_d=w(t)\,dt \qquad d\mathbf{w}_p=\mu(t)\,dt. \tag{42}$$

The stationary probability density functions of states of the dominant part of equation (40) are independent Gaussian densities as

$$p_d^t(x_1)=\frac{1}{(2\pi E_d[x_1^2])^{1/2}}\exp\left(-\frac{x_1^2}{2E_d[x_1^2]}\right) \tag{43}$$

$$p_d^t(x_2)=\frac{1}{(2\pi E_d[x_2^2])^{1/2}}\exp\left(-\frac{x_2^2}{2E_d[x_2^2]}\right) \tag{44}$$

where

$$E_d[x_1^2]=\frac{q_0}{2\xi_0\omega_0^3} \tag{45}$$

$$E_d[x_2^2]=\frac{E_d[x_2^2]}{\omega_0^2} \tag{46}$$

Substituting $\phi(\mathbf{x})=x_1^{r_1}x_2^{r_2}$ and equation (42) into (20), and utilizing the property that x_1 and x_2 are independent in evaluating $E_d[\cdot]$, one derives

$$\sum_{l_1=0}^{m_1}\sum_{l_2=0}^{m_2}\lambda^p_{l_1l_2 0..0}\{r_1E_d[x_1^{r_1+l_1-1}]\,E_d\,[x_2^{r_2+l_2+1}]$$

except all $l_i=0$.

$$-2\xi_0\omega_0 r_2E_d\,[x_1^{r_1+l_1}]\,E_d\,[x_2^{r_2+l_2}]-\omega_0^2r_2E_d[x_1^{r_1+l_1+1}]E_d[x_2^{r_2+l_2-1}] + q_0r_2(r_2-1)E_d[x_1^{r_1+l_1}]E_d[x_2^{r_2+l_2-2}]\} + q_1r_2(r_2-1)E_d\,[x_1^{r_1+2}]E_d[x_2^{r_2-2}]=0. \tag{47}$$

By choosing $m_1=m_2=2$ and noting that the probability density functions of states in equation (43) and (44) are even functions, the following equations can be derived as

$r_1=1$, $r_2=0$:

$$\lambda^p_{010..0}=0 \tag{48}$$

$r_1=0$, $r_2=1$:

$$\lambda^p_{100..0}=0 \tag{49}$$

$r_1=2$, $r_2=0$:

$$\lambda^p_{110..0}=0 \tag{50}$$

$r_1=1$, $r_2=1$:

$$(E_d[x_1^2]E_d[x_2^2]-\omega_0^2E_d[x_1^4])\lambda^p_{200..0} + (E_d[x_2^4]-\omega_0^2E_d[x_1^2]E_d[x_2^2])\lambda^p_{020..0}=0 \tag{51}$$

$r_1=0$, $r_2=2$:

$$(-2\xi_0\omega_0E_d[x_1^2]E_d[x_2^2]+q_0E_d[x_1^2])\lambda^p_{200..0} + (-2\xi_0\omega_0E_d[x_2^4]+q_0E_d[x_2^2])\lambda^p_{020..0}+q_1E_d[x_1^2]=0. \tag{52}$$

From equation (51) and employing (43) to (46), one has

$$\lambda^p_{200..0}=\omega_0^2\lambda_{020..0}. \tag{53}$$

From equations (52), (53) and utilizing (43) to (46), one obtains

$$\lambda^p_{020..0}=\frac{q_1}{2q_0\omega_0^2}. \tag{54}$$

Thus, the stationary probability density functions of x_1 and x_2 can be expressed as equations (22) by substituting the derived $\lambda^p_{ij0..0}$ in the above equations to yield

$$p^t(x_1)=\frac{1}{(\pi/k)^{1/2}}\exp(-kx_1^2) \tag{55}$$

$$p^t(x_2)=\frac{1}{(\pi\omega_0^2/k)^{1/2}}\exp\left(-\frac{kx_2^2}{\omega_0^2}\right) \tag{56}$$

where

$$k=\frac{2\xi_0\omega_0^3-\epsilon q_1}{2q_0}>0. \tag{57}$$

The validity of the derived probability density function is supported by estimating the exact density function (Bendat and Piersol, 1971) through Monte Carlo simulations. By choosing $\xi_0=0.5$, $\omega_0=1.0$, $q_0=0.5$, $q_1=0.5$ with $\epsilon=0.1$ and 0.5, respectively, the estimated probability density functions of x_1 are compared with the derived density function (55) as shown in Figs. 1 and 2.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

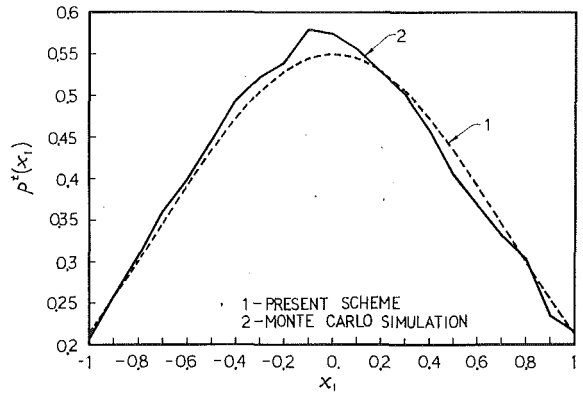


Fig. 1 Stationary probability density function of x_1 of a linear oscillator with a linear spring and noise multiplicative term, $\epsilon = 0.1$ (i.e., $\epsilon^{1/2} = 0.316$)

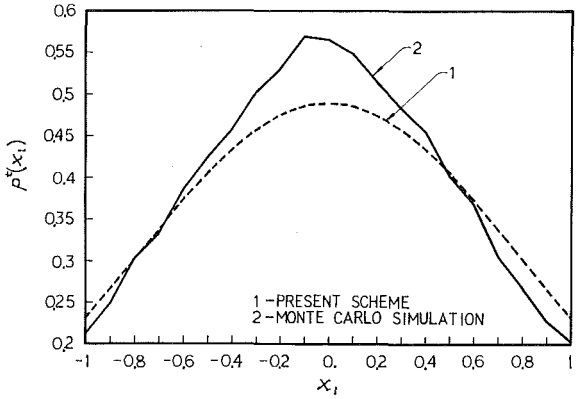


Fig. 2 Stationary probability density function of x_1 of a linear oscillator with a linear spring and noise multiplicative term, $\epsilon = 0.5$ (i.e., $\epsilon^{1/2} = 0.707$)

From Fig. 1, it is observed that equation (55) can be an approximate stationary density function of equation (40) when ϵ is equal to 0.1 (i.e., $\epsilon^{1/2} = 0.316$). When ϵ is increased to 0.5 (i.e., $\epsilon^{1/2} = 0.707$), the inaccuracy of using equation (55) for the stationary density function of (40) can be seen as shown in Fig. 2. If $\epsilon^{1/2}$ is less than 0.1, accurate approximation of the stationary density function of (40) by (55) can be obtained as expected.

In this example, the present technique is applied to derive the stationary probability density function of states of a linear oscillator including state-noise multiplicative term. The final example selected for illustration is a nonlinear oscillator under both parametric and external noise excitations.

Example 3. Consider a second-order nonlinear stochastic oscillator governed by

$$\ddot{x} + (2\xi_0\omega_0 + \xi(t))\dot{x} + (\omega_0^2 + \mu(t))x + \epsilon x^3 = w(t) \tag{58}$$

where $\xi(t)$, $\mu(t)$, and $w(t)$ are mutually independent physical Gaussian white noises with intensities 2^{q_2}, 2^{q_1}, and 2^{q_0}, respectively.

The nonlinear stochastic oscillator can be expressed as equation (1) in the Ito sense with the system parameters

$$\mathbf{f}_d = \begin{pmatrix} x_2 \\ -(2\xi_0\omega_0 - q_2)x_2 - \omega_0^2 x_1 \end{pmatrix} \quad d\mathbf{w}_d = \begin{pmatrix} \xi(t) \\ \mu(t) \\ w(t) \end{pmatrix} dt \quad d\mathbf{w}_p = 0$$

$$G_d = \begin{pmatrix} 0 & 0 & 0 \\ -x_2 & -x_1 & 1 \end{pmatrix} \quad \mathbf{f}_p = \begin{pmatrix} 0 \\ -x_1^3 \end{pmatrix}. \tag{59}$$

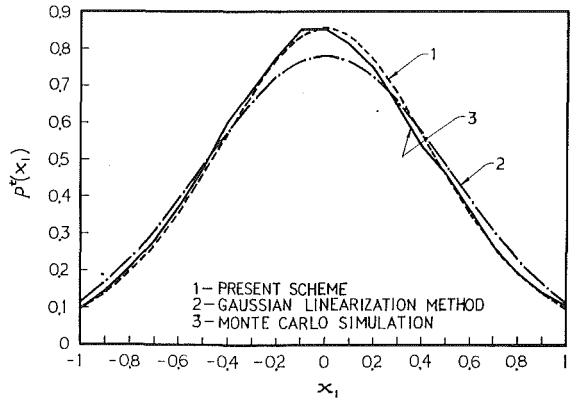


Fig. 3 Stationary probability density function of x_1 of a parametric and external noise-excited Duffing-type oscillator, $\epsilon = 0.1$

The stationary probability density functions of states of equation (58) with $\epsilon = 0$ have been derived by Dimentberg (1982) for $q_1 = \omega_0^2 \cdot q_2$ as

$$p_d^t(\mathbf{x}_1) = \frac{\alpha^{\beta-1}\Gamma\left(\beta - \frac{1}{2}\right)}{\pi^{1/2}\Gamma(\beta-1)} (\alpha + x_1^2)^{-(\beta-1/2)} \tag{60}$$

and

$$p_d^t\left(\frac{x_2}{\omega_0}\right) = p_d^t(x_1) \tag{61}$$

where

$$\alpha = \frac{q_0}{q_1} \tag{62}$$

$$\beta = \left(\frac{2\xi_0\omega_0}{q_2} + 1\right)/2 \tag{63}$$

By selecting $m_1 = m_2 = 2$ and following the same procedure as before except that x_1 and x_2 are not independent again, the unknown parameters $\lambda_{ij0..0}^p$ in equation (22) can be derived to yield

$$\lambda_{100..0}^p = \lambda_{010..0}^p = \lambda_{110..0}^p = \lambda_{020..0}^p = 0$$

$$\lambda_{20..0} = -\frac{1.5}{\omega_0^2}. \tag{64}$$

Thus, the approximate probability density function of x_1 can be expressed by substituting the derived $\lambda_{ij0..0}^p$ and equations (60) into equation (22) to give

$$p^t(x_1) = N_1(\alpha + x_1^2)^{-(\beta-1/2)} \exp\left(-\frac{1.5\epsilon}{\omega_0^2} x_1^2\right) \tag{65}$$

where N_1 is a normalization constant.

The validity of the approximate stationary probability density function of x_1 derived as equation (65) will be supported by the estimated density function through Monte Carlo simulation. By selecting $\xi_0 = 0.5$, $\omega_0 = 1.0$, $q_1 = 0.1$, $q_0 = 0.2$, and $\epsilon = 0.1$, the simulated probability density functions by employing equation (65), the Gaussian linearization approach (Chang and Young, 1988), and Monte Carlo method are shown in Fig. 3. For this example, it is seen that the approximate stationary probability density function of x_1 derived as equation (65) is very accurate when ϵ is less than 0.1.

5 Conclusions

A technique based on the maximum entropy approach for deriving the stationary probability density function of a nonlinear stochastic system has been presented in this paper. By

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

employing the present technique, a minimally prejudiced probability density function can be derived by solving a set of nonlinear algebraic equations. For a wide class of nonlinear stochastic system including perturbed nonlinear functions of states and parametric noise-excited terms, an explicit approximate probability density function can be derived from extending a given exact probability density function of the corresponding dominant stochastic system by the present scheme. The illustrated examples show that the present scheme may provide exact density function for some stochastic oscillators and accurate expression of the stationary density function can be assured if the order of perturbed part is less than one tenth of the corresponding dominant part in a nonlinear stochastic oscillator.

References

Bendat, J. S., and Piersol, A. G., 1971, *Random Data: Analysis and Measurement Procedures*, John Wiley and Sons, New York.

Burg, J. P., 1975, "Maximum Entropy Spectral Analysis," Thesis, Stanford University, Stanford, Calif.

Caughey, T. K., 1986, "On the Response of Non-linear Oscillators to Stochastic Excitation," *Probabilistic Engineering Mechanics*, Vol. 1, No. 1, pp. 2-4.

Cai, G. Q., and Lin, Y. K., 1988, "On Exact Stationary Solutions of Equivalent Non-Linear Stochastic Systems," *Int. J. Non-Linear Mechanics*, Vol. 23, No. 4, pp. 315-325.

Chang, R. J., and Young, G. E., 1988, "Methods and Gaussian Criterion for Statistical Linearization of Stochastic Parametrically and Externally Excited Nonlinear Systems," ASME Winter Annual Meeting, 88-WA/DSC-9, to appear in ASME JOURNAL OF APPLIED MECHANICS.

Crandall, S. H., 1985, "Non-Gaussian Closure Techniques for Stationary Random Vibration," *International Journal of Non-Linear Mechanics*, Vol. 20, No. 1, pp. 1-8.

Dimentberg, M. F., 1982, "An Exact Solution to a Certain Non-Linear Random Vibration Problem," *International Journal of Non-Linear Mechanics*, Vol. 17, No. 4, pp. 231-236.

Ibrahim, R. A., 1985, *Parametric Random Vibration*, Research Studies Press, New York.

Ibrahim, R. A., and Heo, H., 1987, "Stochastic Response of Nonlinear Structures With Parameter Random Fluctuations," *AIAA Journal*, Vol. 25, No. 2, pp. 331-338.

Lin, Y. K., 1970, "First-Excursion Failure of Randomly Excited Structures: II," *AIAA Journal*, Vol. 8, No. 10, pp. 1888-1890.

Liu, Q., and Huw, G. D., 1988, "Application of Non-Gaussian Closure to the Nonstationary Response of a Duffing Oscillator," *International Journal of Non-Linear Mechanics*, Vol. 23, No. 3, pp. 241-250.

Robert, J. B., and Spanos, P. D., 1986, "Stochastic Averaging: An Approximate Method of Solving Random Vibration Problems," *International Journal of Non-Linear Mechanics*, Vol. 21, No. 2, pp. 111-134.

Siddall, J. N., and Diab, Y., 1975, "The Use in Probabilistic Design of Probability Curves Generated by Maximizing the Shannon Entropy Function Constrained by Moments," ASME *Journal of Engineering for Industry*, pp. 843-852.

Tribus, M., 1969, *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.

Young, G. E., and Chang, R. J., 1987, "Prediction of the Response of Non-Linear Oscillators Under Stochastic Parametric and External Excitations," *International Journal of Non-Linear Mechanics*, Vol. 22, No. 2, pp. 151-160.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Ilter Serbetci
NASA Lewis Research Center,
Space Propulsion Technology Division,
Cleveland, OH 44735
Mem. ASME

John A. Tichy
Mem. ASME

Department of Mechanical Engineering,
Aeronautical Engineering and Mechanics,
Rensselaer Polytechnic Institute,
Troy, NY 12181

Inertia Effects in Thin Film Flow With a Corrugated Boundary

An analytical solution is presented for two-dimensional, incompressible film flow between a sinusoidally grooved (or rough) surface and a flat surface. The upper grooved surface is stationary whereas the lower, smooth surface moves with a constant speed. The Navier-Stokes equations were solved employing both mapping techniques and perturbation expansions. Due to the inclusion of the inertia effects, a different pressure distribution is obtained than predicted by the classical lubrication theory. In particular, the amplitude of the pressure distribution of the classical lubrication theory is found to be in error by over 100 percent (for modified Reynolds number of 3–4).

Introduction

The behavior of thin, incompressible Newtonian fluid films between grooved (or rough) surfaces has been studied by many investigators. This behavior is a developing Couette-type flow which incorporates both a pressure gradient (Poiseuille) and a no-slip boundary condition as the driving mechanisms. This type of a flow is not only of academic interest in fluid mechanics but also of interest in practical tribological applications, for it describes the lubricant flow in slider bearings with grooved (or rough) surfaces. Because of the rough surfaces flow reversals occur within the film and the resulting streamlines are not straight. The grooves (or roughness) may be intentional to increase the load capacity and to inject lubricant or may be the result of imprecise machining and wear.

One of the earliest studies of this phenomenon is due to Michell (1950) who studied the effect of sinusoidal longitudinal roughness on slider bearings for a Newtonian-constant property fluid. Elrod (1973) also analyzed the effects of oblique sinusoidal roughness on the pressure distribution starting with the Reynolds equation. For both longitudinal and transverse roughness, Christensen and Tonder (1971) attacked the problem with an averaged Reynolds equation for finite-width inclined bearings using a probability density function for roughness.

Wang (1976, 1978) studied Couette flow between corrugated plates. His analysis showed that both the flow rate and the wall shear stress are functions of the phase shift between sinusoidally grooved plates. A regular perturbation expansion was used for the Navier-Stokes equations to predict the velocity distributions. Inertia terms were ignored in this analysis. Sun and Cheng (1977) analyzed roughness effects on hydrodynamic lubrication using Stokes equations. While shear terms were included, inertia effects were again neglected in this analysis. White (1980) also analyzed thin compressible lubrication films. For an infinite-width parallel slider bearing, he showed that the one-dimensional load capacity varies with ϵ, where ϵ is the roughness amplitude to average film thickness ratio. Even though this functional dependence is not valid for values of ϵ close to 1, it demonstrates the tendency of the load capacity to increase with roughness amplitude. White obtained these results using an averaged Reynolds equation. Similarly, Tonder (1977) used three different averaged Reynolds equation for transverse, longitudinal, and isotropic area-distributed roughness.

In the works previously cited, either the Reynolds equation or an averaged Reynolds equation was used. The Reynolds equation is a one-dimensional version of the three-dimensional momentum equations with no inertia terms. Exclusion of the inertia terms can be justified in the case of small modified Reynolds numbers (i.e., $\mathrm{Re}^* = Uh_o^2/\nu\lambda << 1$; where U is the speed of the moving plate, h_o is the average film thickness, λ is the average wavelength of roughness, and ν is the viscosity). However, in modern applications, increased machine speeds and low viscosity lubricants may preclude one from neglecting inertia terms. In fact, for high modified Reynolds numbers (i.e., $\mathrm{Re}^* \geq 3$), inertia terms start to have a large effect on the pressure and velocity distributions.

Most of the recent works in which inertia terms are included are concerned with squeeze film flow between oscillating surfaces (Tichy and Winer (1970), Tichy and Modest (1978), Grimm (1976)) rather than with roughness in Couette-type flows. In one recent work, Hasegawa and Izuchi (1984) included inertia terms in their analysis of channel flow between an uneven and a plane wall. This work concentrates on nonperiodic portions of the tertiary components of drag and mass flow rate with the perturbation expansions being based on small wave numbers rather than small undulation amplitudes. However, the periodic secondary component of the stream function determines the amplitude of the pressure distribution even

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 2, 1989; final revision, May 15, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

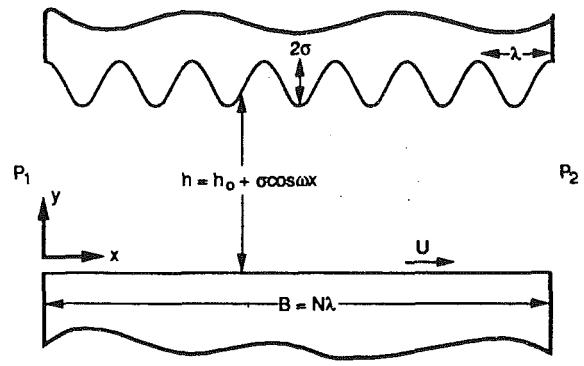


Fig. 1 Thin film channel flow with a rough boundary

though it does not affect either drag or mass flow rate magnitude. The present work contributes to the understanding of this phenomenon. In order to account for the inertia effects, convective terms of the Navier-Stokes equations are retained and can be linearized due to the small amplitude of disturbance induced by the rough boundary. Since the inertia terms are not neglected, the results are valid for any Re* which enables us to observe interesting effects not predicted by lubrication theory.

Analysis

The thin film geometry is depicted in Fig. 1, using an exaggerated length scale in the vertical direction. The upper surface is sinusoidally corrugated and the film thickness is given by $h = h_o + \delta \cos(\omega x)$ where h_o represents the average film thickness. Peak-to-peak roughness amplitude is designated 2δ and ω is the roughness frequency. The bottom surface is assumed to be smooth and moving with constant speed U, whereas the upper surface is stationary. Pressure boundary conditions are p_1 and p_2 at the inlet and outlet, respectively. Relative sizes are as follows,

$$h_o/\lambda << 1,\ N >> 1,\ L >> B,\ B = N\lambda$$

where N is the number of grooves and L represents the length in the z direction.

Classical Lubrication Solution. Classical lubrication approach makes use of the Reynolds equation which is derived specifically for thin film inertialess flow from the Navier-Stokes equations. For a steady, long channel the Reynolds equation reads

$$\frac{d}{dx}\left(h^3 \frac{dp}{dx}\right) = 12\mu \frac{d}{dx}\left(\frac{hU}{2}\right). \tag{1}$$

Using the nondimensional quantities,

$$p^* = \frac{p - p_1}{U\mu\lambda/h_o^2},\ \xi = x/\lambda,\ \epsilon = \delta/h_o,\ h^* = h/h_o = 1 + \epsilon \cos 2\pi\xi,$$

the equation is given the form

$$\frac{d}{d\xi}\left(h^{*3} \frac{dp^*}{d\xi}\right) = 6\frac{dh^*}{d\xi} \tag{2}$$

with boundary conditions

$$p^* = p_1^* = 0 \text{ at } \xi = 0$$
$$p^* = p_2^* \text{ at } \xi = N.$$

The solution to equation (2) is obtained by successive integrations. The first integration yields

$$\frac{dp^*}{d\xi} = 6\left(\frac{1}{h^{*2}} + \frac{c}{h^{*3}}\right), \tag{3}$$

where c is the integration constant to be determined from the boundary condition

$$p^*(\xi = N) = p_2^*,$$
$$p_2^* = 6\int_0^N \frac{d\xi}{h^{*2}} + 6\int_0^N \frac{c}{h^{*3}}\, d\xi. \tag{4}$$

Using Booker's integral tables (1965) one obtains

$$c = \frac{\dfrac{2p_2^*(1-\epsilon^2)^{3/2}}{6N} - 2}{\dfrac{3}{(1-\epsilon^2)} - 1}. \tag{5}$$

Employing Taylor expansion series, the terms in the denominator of equation (3) becomes

$$\frac{1}{h^{*2}} = (1 + \epsilon \cos 2\pi\xi)^{-2} \approx 1 - \frac{2!}{1!}\epsilon \cos(2\pi\xi) + \frac{3!}{2!}\epsilon^2 \cos^2(2\pi\xi) - \ldots \tag{6}$$

$$\frac{1}{h^{*3}} = (1 + \epsilon \cos 2\pi\xi)^{-3} \approx 1 - \frac{3!}{2!}\epsilon \cos(2\pi\xi) + \frac{4!}{2!2!}\epsilon^2 \cos^2(2\pi\xi) - \ldots . \tag{7}$$

Now, equation (3) can be integrated to yield the pressure distribution in the film,

$$p^*(\xi) = 6\left\{\xi - \frac{2\epsilon}{2\pi}\sin 2\pi\xi + 3\epsilon^2\left(\frac{\xi}{2} + \frac{\sin 4\pi\xi}{4\pi}\right)\right\} + 6\left\{\xi - \frac{3\epsilon}{2\pi}\sin 2\pi\xi + 6\epsilon^2\left(\frac{\xi}{2} + \frac{\sin 4\pi\xi}{4\pi}\right)\right\}\frac{\dfrac{2p_2^*(1-\epsilon^2)^{3/2}}{6N} - 2}{\dfrac{3}{(1-\epsilon^2)} - 1}. \tag{8}$$

Note that for $\epsilon << 1$, the pressure distribution is sinusoidal and -90 deg out of phase with the roughness distribution;

$$p^* = p^o + \epsilon p' \sin 2\pi\xi + \epsilon^2 \ldots = p^o + \epsilon p' \cos(2\pi\xi - 90°) + \epsilon^2 \ldots \tag{9}$$

where

$$p^o = \frac{\xi}{N}p_2^* \text{ and } p' = \frac{3}{\pi}\left(1 - \frac{p_2^*}{2N}\right).$$

Present Study With Approximated Inertia Terms. The continuity and full momentum equations for the incompressible Newtonian film under consideration appears as

$$\left.\begin{aligned} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= 0 \\ u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \\ u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial y} + \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) \end{aligned}\right\} \tag{10}$$

with the *known* velocity boundary conditions

$$\left.\begin{aligned} u(x,h) &= 0 \quad & v(x,h) &= 0 \\ u(x,0) &= U \quad & v(x,0) &= 0 \end{aligned}\right\}. \tag{11}$$

These equations can be tailored for the given geometry by performing an order-of-magnitude analysis. Since

$\delta << h_o << \lambda$, $\partial/\partial x \sim 1/\lambda$ and $\partial/\partial y \sim 1/h_o$ and $v \sim Uh_o/\lambda$,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

final equations reduce to boundary layer equations

$$\left.\begin{aligned} u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}&=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\frac{\partial^2 u}{\partial y^2}\\ 0&\cong-\frac{1}{\rho}\frac{\partial p}{\partial y}\end{aligned}\right\}. \tag{12}$$

This indicates that pressure variation across the film is negligible and $p = p(x)$ only. Therefore, the known boundary conditions are sufficient to solve the problem.

Introducing a nondimensional stream function, $u = Uh_o\partial\psi/\partial y$, $v = -Uh_o\partial\psi/\partial x$, the continuity equation is satisfied. Defining nondimensional variables, $\xi = x/\lambda$, $\eta = y/h(x)$ and cross-differentiating equation (12), the pressure gradients are eliminated and a fourth-order nonlinear variable coefficient, partial differential equation is developed:

$$\frac{h}{h_o}\frac{\partial\psi}{\partial\eta}\frac{\partial^3\psi}{\partial\xi\partial\eta^2}+2\frac{\partial\psi}{\partial\eta}2\pi\epsilon\sin(2\pi\xi)\frac{\partial^2\psi}{\partial\eta^2}+\frac{h}{h_o}\frac{\partial\psi}{\partial\xi}\frac{\partial^3\psi}{\partial\eta^3}=\frac{1}{\mathrm{Re}^*}\frac{\partial^4\psi}{\partial\eta^4}. \tag{13}$$

If there were no roughness, the stream function would be a function of η only. However, the presence of upper surface roughness induces a functional variation in the ξ direction as well. For small values of ϵ, the stream function can be expressed by a regular perturbation expansion in ϵ,

$$\psi(\eta,\xi,\epsilon)=\psi^o(\eta)+\epsilon\psi'(\eta,\xi)+\epsilon^2\psi''(\eta,\xi)+\ldots. \tag{14}$$

Substitution of this power expansion into equation (13) and using $h/h_o = 1 + \epsilon\cos 2\pi\xi$ yields differential equations for powers of ϵ:

$$(O)\epsilon^0:\ \psi^o_{\eta\eta\eta\eta}=0 \tag{15}$$

$$(O)\epsilon^1:\ \psi^o\psi'_{\eta\eta\xi}+4\pi\sin(2\pi\xi)\psi^o_\eta\psi^o_{\eta\eta}-\psi_\xi\psi^o_{\eta\eta\eta}=\frac{1}{\mathrm{Re}^*}\psi'_{\eta\eta\eta\eta} \tag{16}$$

$$\begin{aligned}(O)\epsilon^2:\ &\psi^o_\eta\psi'_{\eta\eta\xi}\cos(2\pi\xi)+\psi^o_\eta\psi''_{\eta\eta\xi}+4\pi\psi^o_\eta\psi'_{\eta\eta}\sin(2\pi\xi)\\ &+4\pi\psi'_\eta\psi^o_{\eta\eta}\sin(2\pi\xi)-\psi'_{\eta\eta\eta}\psi_\xi-\psi^o_{\eta\eta\eta}\psi_\xi\cos(2\pi\xi)\\ &-\psi^o_{\eta\eta\eta}\psi_\xi''=\psi''_{\eta\eta\eta\eta}/\mathrm{Re}^*\end{aligned} \tag{17}$$

$$(O)\epsilon^3:\ \ldots$$

which are subject to the nondimensional boundary conditions

$$\left.\frac{\partial\psi}{\partial\eta}\right|_{\eta=1}=0 \tag{18}$$

$$\left.\frac{\partial\psi}{\partial\xi}\right|_{\eta=1}=0 \text{ or } \psi(\eta=1)=Q \tag{19}$$

$$\left.\frac{\partial\psi}{\partial\eta}\right|_{\eta=0}=(1+\epsilon\cos 2\pi\xi) \tag{20}$$

$$\left.\frac{\partial\psi}{\partial\xi}\right|_{\eta=0}=0 \text{ or } \psi(\eta=0)=C. \tag{21}$$

The choice of one of the constants, C or Q, is arbitrary. Since the difference between any two stream functions gives the mass flow rate enclosed by those stream functions, C is chosen to be zero, letting Q be the mass flow rate,

$$Q=\psi(\xi,1)-\psi(\xi,0).$$

The general solution to equation (15) is given by:

$$\psi^o(\eta)=A\eta^3+B\eta^2+C\eta+D. \tag{22}$$

Making use of the boundary conditions,

$$@\eta=0,\ \psi^o=0,\ \psi^o_\eta=1;\ @\eta=1,\ \psi^o=Q,\ \psi^o_\eta=0$$

the constants in equation (22) are determined so that,

$$\psi^o(\eta)=(1-2Q)\eta^3+(3Q-2)\eta^2+\eta. \tag{23}$$

Upon substituting ψ^o_η, $\psi^o_{\eta\eta}$, and $\psi^o_{\eta\eta\eta}$ into equation (16), the resulting equation can be expressed in terms of ψ' (or its derivatives) only

$$\alpha(\eta)\psi'_{\eta\eta\xi}-\beta\psi'_\xi+\gamma(\eta)2\pi\sin(2\pi\xi)=\psi'_{\eta\eta\eta\eta}/\mathrm{Re}^* \tag{24}$$

where

$$\alpha(\eta)=3(1-2Q)\eta^2+2(3Q-2)\eta+1$$

$$\beta=6(1-2Q)$$

$$\gamma(\eta)=36(1-Q)^2\eta^3+36(3Q-2)(1-2Q)\eta^2+12(1-2Q)\eta+8(3Q-2)^2+4(3Q-2).$$

Since the upper boundary condition imposes a sinusoidal stream function and the perturbation equations are linear, it is suitable to designate the stream function in a complex form. In this way ξ-dependence may be temporarily eliminated and equation (24) can be transformed into a set of ordinary differential equations. The complex form of ψ' is defined as

$$\psi'=\mathrm{Real}\,[\bar\psi'(\eta)e^{i2\pi\xi}]=\mathrm{Real}\,[\bar\psi']$$

where

$$\bar\psi'=\bar\psi'^r+i\bar\psi'^i.$$

Equation (24) can be rearranged to read

$$\alpha(\eta)i2\pi\bar\psi'_{\eta\eta}e^{i2\pi\xi}-\beta i2\pi\bar\psi'e^{i2\pi\xi}-\gamma(\eta)2\pi ie^{i2\pi\xi}=1/\mathrm{Re}^*\bar\psi'_{\eta\eta\eta\eta}e^{i2\pi\xi} \tag{25}$$

in which $\sin 2\pi\xi$ is treated as Real $[-ie^{i2\pi\xi}]$. The real and imaginary parts of this equation constitute two coupled, fourth-order "quasi-linear," variable-coefficient ordinary differential equations,

$$\alpha(\eta)2\pi\bar\psi'^r_{\eta\eta}-\beta2\pi\bar\psi'^r_\eta-\gamma(\eta)2\pi=1/\mathrm{Re}^*\bar\psi'^i_{\eta\eta\eta\eta} \tag{26}$$

$$-\alpha(\eta)2\pi\bar\psi'^i_{\eta\eta}+\beta2\pi\bar\psi'^i_\eta=1/\mathrm{Re}^*\bar\psi'^i_{\eta\eta\eta\eta} \tag{27}$$

Equations (26), (27) were solved simultaneously for $\bar\psi'(\eta,\xi)$ by numerical means (see Appendix).

Knowing both $\psi^o(\eta)$ and $\bar\psi'(\eta,\xi)$ one can obtain $\bar\psi''(\eta,\xi)$ from equation (17) in a similar way.

Secondary velocity components, $\bar\psi'^r$ and $\bar\psi'^i$ are depicted in Fig. 2 for $p_2^* = -900$, $Q = 2$, $h_o/\lambda = 0.1$ and for various modified Reynolds numbers. When $\mathrm{Re}^* = 0.0$, $\bar\psi'^i$ is identically zero everywhere, and $\bar\psi'^r$ displays a parabolic profile to meet the boundary condition (20). This case represents the classical lubrication solution in the transferred coordinate system. As Re^* is increased both $\bar\psi'^r$ and $\bar\psi'^i$ start to contribute to total velocity with increasing amplitude. The *positive* and *negative* areas in the figures indicate flow reversals, but, as expected, there is no contribution to the mass flow rate by either secondary velocity component as they cancel out.

Pressure Distribution. Starting with the tailored momentum equation of the previous part, equation (12), the pressure gradient can be expressed in terms of ψ as

$$\begin{aligned}\frac{dp^*}{d\xi}=-\left(\frac{h_o}{h}\right)^2\Bigg\{&\mathrm{Re}^*\frac{\partial\psi}{\partial\eta}\left[\frac{\partial^2\psi}{\partial\eta\partial\xi}+\frac{h_o}{h}\eta2\pi\epsilon\sin(2\pi\xi)\frac{\partial^2\psi}{\partial\eta^2}\right.\\ &\left.+\frac{h_o}{h}2\pi\epsilon\sin(2\pi\xi)\frac{\partial\psi}{\partial\eta}\right]-\mathrm{Re}^*\frac{\partial^2\psi}{\partial\eta^2}\left[\frac{\partial\psi}{\partial\xi}+\frac{h_o}{h}\eta2\pi\epsilon\sin(2\pi\xi)\frac{\partial\psi}{\partial\eta}\right]\\ &-\left(\frac{h_o}{h}\right)\frac{\partial^3\psi}{\partial\eta^3}\Bigg\}.\end{aligned} \tag{28}$$

The powers of h/h_o can again be expanded using equations (6) and (7). Recalling that $\psi(\eta,\xi,\epsilon) = \psi^o(\eta) + \epsilon\psi'(\eta,\xi) + \epsilon^2\psi''(\eta,\xi)+\ldots$, equation (28) is arranged to yield differential equations of different orders of magnitude:

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$(O)\epsilon^0 : -\frac{\partial p^o}{\partial \xi} = -\psi^o_{\eta\eta\eta} = -6(1-2Q) \tag{29}$$

$$(O)\epsilon^1 : -\frac{\partial p'}{\partial \xi} = \mathrm{Re}^*(\psi'_{\eta\xi}\psi^o_\eta + 2\pi \sin(2\pi\xi)\psi^{o2}_\eta - \psi'_\xi\psi^o_{\eta\eta}) - \psi'_{\eta\eta\eta} + 3\cos(2\pi\xi)\psi^o_{\eta\eta\eta} \tag{30}$$

$$(O)\epsilon^2 : -\frac{\partial p''}{\partial \xi} = \mathrm{Re}^*\{\psi'_\eta\psi'_{\eta\xi} + \psi^o_\eta\psi''_{\xi\eta} - 2\psi^o_\eta\cos(2\pi\xi)\psi'_{\xi\eta} + \psi^o_\eta 2\pi \sin(4\pi\xi)\psi'_\eta + \psi'_\eta 2\pi \sin(4\pi\xi)\psi^o_\eta - 6\pi\psi^{o2}_\eta\cos(2\pi\xi)\sin(2\pi\xi) - \psi^o_{\eta\eta}\psi''_\xi + 2^o_{\eta\eta}\psi'_\xi\cos(2\pi\xi) - \psi'_{\eta\eta}\psi'_\xi\} - (\psi''_{\eta\eta\eta} - 3\psi'_{\eta\eta\eta}\cos(2\pi\xi) + 6\psi^o_{\eta\eta\eta}(\cos^2(2\pi\xi)) \tag{31}$$

Upon substituting

$$\psi' = \bar{\psi}'^r\cos(2\pi\xi) - \bar{\psi}'^i\sin(2\pi\xi)$$

$$\psi'' = \theta''(\eta) + \bar{\psi}''^r\cos(4\pi\xi) - \bar{\psi}''^i\sin(4\pi\xi),$$

the pressure gradient for the fluid film is given by

$$\frac{dp^*}{d\xi} = K_1 + \epsilon\mathrm{Re}^*\{K_2\sin(2\pi\xi) + K_3\cos(2\pi\xi)\} + \epsilon^2\mathrm{Re}^*\left\{\frac{\theta''_{\eta\eta\eta}}{\mathrm{Re}^*} + K_4\sin(2\pi\xi) + K_5\cos(2\pi\xi) + K_6\cos(2\pi\xi)\sin(2\pi\xi) + K_7\cos^2(2\pi\xi) + K_8\sin^2(2\pi\xi)\right\} \tag{32}$$

where

$$K_1 = 6(1-2Q)$$

$$K_2 = 2\pi\alpha(\eta)\bar{\psi}'^r_\eta - 2\pi\alpha^2(\eta) - 2\pi Z(\eta)\bar{\psi}'^r - \bar{\psi}'^i_{\eta\eta\eta}/\mathrm{Re}^*$$

$$K_3 = 2\pi\alpha(\eta)\bar{\psi}'^i_\eta - 2\pi Z(\eta)\bar{\psi}'^i - (3K_1 - \bar{\psi}'^r_{\eta\eta\eta})/\mathrm{Re}^*$$

$$K_4 = 4\pi\alpha(\eta)\bar{\psi}''^r_\eta + 4\pi Z(\eta)\bar{\psi}''^r - \bar{\psi}''^i_{\eta\eta\eta}/\mathrm{Re}^*$$

$$K_5 = 4\pi\alpha(\eta)\bar{\psi}''^i_\eta - 4\pi Z(\eta)\bar{\psi}''^i + \bar{\psi}''^r_{\eta\eta\eta}/\mathrm{Re}^*$$

$$K_6 = -4\pi\alpha(\eta)\bar{\psi}'^r_\eta + 2\pi\bar{\psi}'^{r2}_\eta - 2\pi\bar{\psi}'^{i2}_\eta - 4\pi\alpha(\eta)\bar{\psi}'^r_\eta + 6\pi\alpha^2(\eta) + 4\pi Z(\eta)\bar{\psi}'^r - 2\pi\bar{\psi}'^r_{\eta\eta}\bar{\psi}'^r + 2\pi\bar{\psi}'^i_{\eta\eta}\bar{\psi}'^i + 3\bar{\psi}'^i_{\eta\eta\eta}/\mathrm{Re}^*$$

$$K_7 = -4\pi\alpha(\eta)\bar{\psi}'^i_\eta + 2\pi\bar{\psi}'^i_\eta\bar{\psi}'^r_\eta + 4\pi Z(\eta)\bar{\psi}'^i - 2\pi\bar{\psi}'^r_{\eta\eta}\bar{\psi}'^i - (2\bar{\psi}'^r_{\eta\eta\eta} - 6K_1)/\mathrm{Re}^*$$

$$K_8 = 2\pi\bar{\psi}'^i_\eta\bar{\psi}'^r + 4\pi\alpha(\eta)\bar{\psi}'^i_\eta + 2\pi\bar{\psi}'^i_{\eta\eta}\bar{\psi}'^r$$

and $Z(\eta) = 6(1-2Q)\eta + 2(3Q-2)$.

The pressure distribution in its final form is obtained by integrating equation (32) and rearranging so that

$$P^*(\xi) = K_1\xi + \epsilon\mathrm{Re}^*\{A_o\cos(2\pi\xi + \phi)\} + \epsilon^2\mathrm{Re}^*. \ldots \tag{33}$$

where

$$A_o = \frac{1}{2\pi}[K_2^2 + K_3^2]^{1/2},\ \phi = \tan^{-1}\{K_3/K_2\} - 180°.$$

Neglecting the contribution of the ϵ^2 term, the pressure phase shift is given by ϕ within a good accuracy.

Discussion and Conclusions

Similar to the classical lubrication solution, equation (9), the present study also predicts a sinusoidal pressure distribution, as given by equation (33). However, the amplitude of the pressure variation changes considerably with Re* as presented in Fig. 3. The nondimensional pressure amplitude predicted by the classical lubrication solution is about 0.95 for $\epsilon = 0.1$, $p^*_2 = -900$ and $N = 50$. As expected, the present study reduces to the classical lubrication solution for Re * << 1. For Re* $= 1 \times 10^{-5}$ and 0.1, the nondimensional pressure amplitudes are 0.95, and 0.96, respectively. This range describes inertialess flow; therefore, the basic assumption of the classical lubrication solution is fulfilled. However, as the modified Reynolds number gets closer to one, the present theory starts diverging from classical lubrication theory. For Re* = 0.7, the error due to inertialess flow assumption is about 25 percent; therefore, a range of $0 < \mathrm{Re}^* < 0.7$ seems to be the operating region for the classical lubrication solution for practical engineering applications. For Re* = 2.0 and 5.0, the nondimensional pressure amplitudes are 2.7 and 6.3, respectively. Therefore, in the range $1.5 < \mathrm{Re}^* < 2.5$, the classical lubrication solution produces results which are in error by more than 100 percent. For even higher modified Reynolds numbers, which were not included in the graph, the nondimensional pressure amplitude increases almost linearly to 8.9 for Re* = 10. As a practical rule, for $\mathrm{Re}^* > 1.5$, the classical lubrication solution should not be employed for pressure calculations since it yields highly erroneous results. Remembering that the pressure distribution is sinusoidal, cavitation problems may result for high modified Reynolds numbers, and such cannot be predicted by the classical lubrication approach. Phase shift between the geometrical roughness and the pressure distribution also appears to be a strong function of Re* (Fig. 3). Again, for Re* << 1, both theories converge to yield a phase shift of -90 deg. As the modified Reynolds number increases, the phase shift tends to -180 deg asymptotically. This is expected from the geometry of the flow passage; as Re* increases, the higher flow inertia becomes more sensitive to surface roughness. One can imagine the undulations as giant asperities acting like hurdles for the flow. These hurdles not only defer the mass flow rate (Wang (1978); Hasegawa and Izuchi (1984)) but also increase or decrease the pressure considerably at decreasing or increasing cross-sectional flow areas, respectively. Negative phase shift indicates that pressure lags roughness.

Ranges for ϵ, h_o/λ *and Re**. It should be remembered that this analysis is based on perturbation expansions in terms of small roughness amplitudes, ϵ. Therefore, one should define a range for the perturbation parameter. For reasonable reliability of the results a range of $0.0 \le \epsilon < 0.3$–0.4 is recommended. Similarly, a practical range for the nondimensional wave number can be set as $h_o/\lambda < 1/10$ to justify the order of magnitude analysis which leads to the boundary layer equations, equation (12). There are no restrictions on the magnitude of the modified Reynolds number providing the flow does not turn turbulent. Since the parallel flow instabilities occur at $\mathrm{Re}^*\lambda/h_o > 2000$ for smooth surfaces, smaller values than this should be adopted for corrugated surfaces, depending on the magnitude of the nondimensional wave number. For small nondimensional wave numbers, e.g., $h_o/\lambda > 1/100$ the effect of roughness on turbulence will be minimal and the parallel flow criterion therefore can be used to determine the nature of the flow, whether laminar or transitional. For turbulent Reynolds numbers one can extend the present work by rewriting equation (12) to include turbulent transport properties and Reynolds' stresses.

For a practical application, consider a turbine oil of 44 universal seconds viscosity ($= 5.5 \times 10^{-6}$ m^2/s) operating at high temperatures (T $\approx$ 200°F). For $U = 15$ m/s, $\lambda = 10$ mm, $h_o = 0.1$ mm, the flow is laminar and Re* is about 2.7. This is outside the acceptable range for the classical lubrication solution to be used; therefore, the present study should be employed. As this example indicates, there are benefits in making a rough calculation first to determine the modified Reynolds number to decide whether or not to use the classical lubrication approach.

In conclusion, (a) for a long channel with a small roughness amplitude, a closed solution to the thin film Navier-Stokes

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

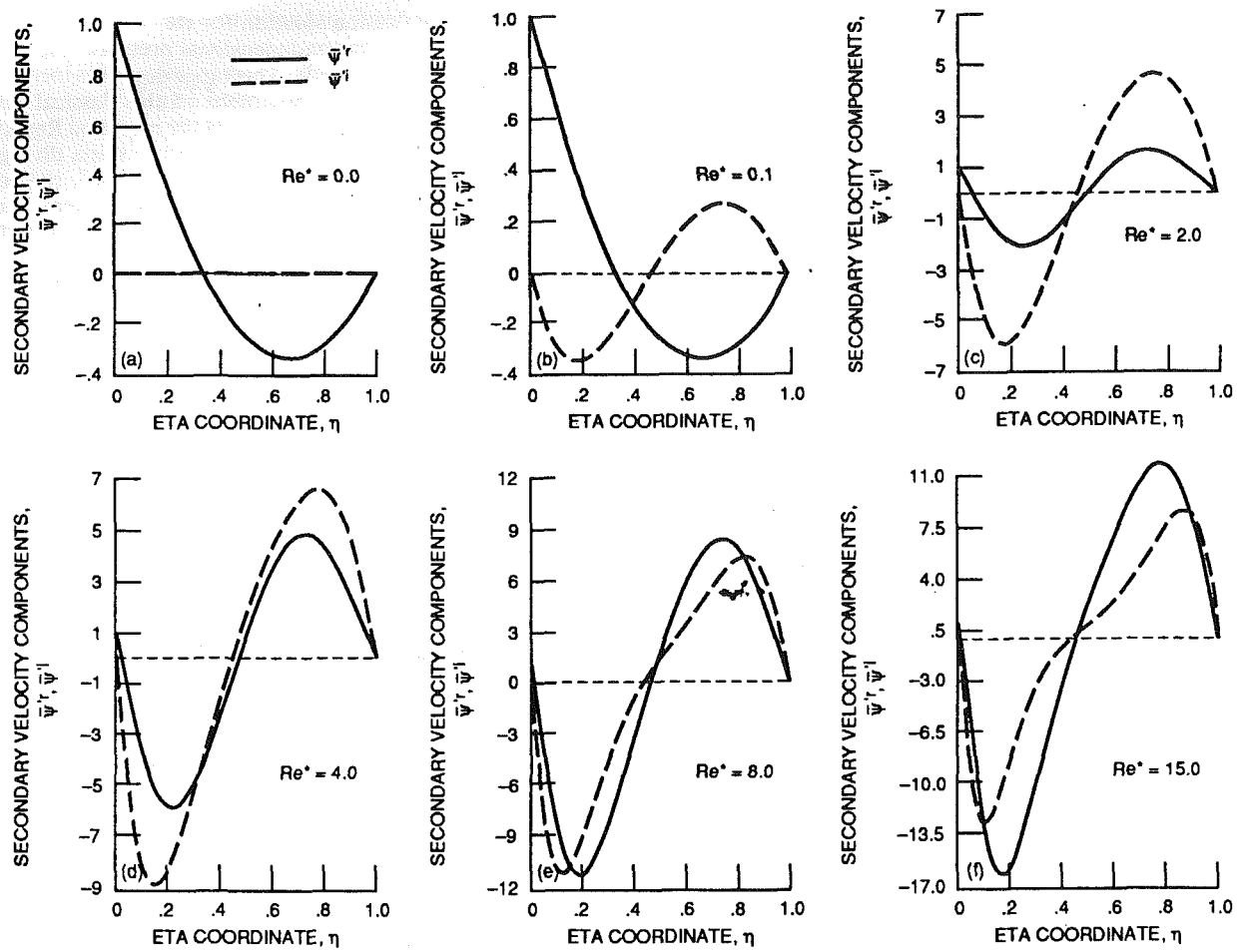


Fig. 2 Effect of Re* on secondary velocity components

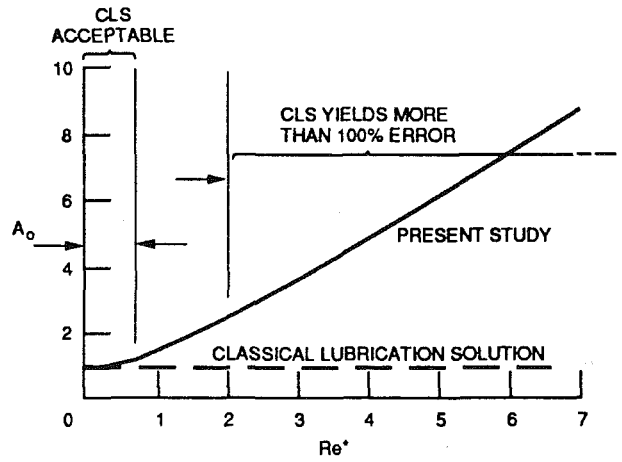


Fig. 3 Effect of Re* on pressure amplitude

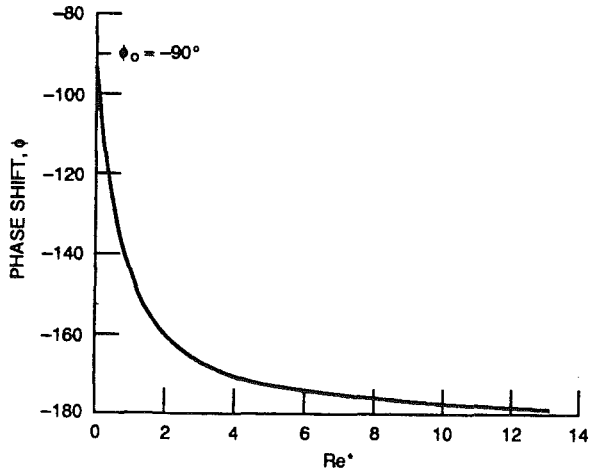


Fig. 4 Effect of Re* on pressure phase shift

equations were obtained accounting for the inertia effects; (b) the pressure amplitude predicted by the classical lubrication solution was found to be in more than 100 percent error for $Re^* \geq 2.0$. The amplitude and the phase shift of the sinusoidal pressure distribution are strong increasing functions of Re* (see Figs. 3 and 4); and (c) the inertia solution reduces to the classical lubrication solution for $Re^* << 1$, as expected.

References

Booker, J. F., 1965, "A Table of the Journal Bearing Integral," *Journal of Basic Engineering*, June, pp. 533–535.

Christensen, H., and Tonder, K., 1971, "The Hydrodynamic Lubrication of Rough Surfaces of Finite Width," ASME *Journal of Lubrication Technology*, July, pp. 234–329.

Elrod, L. H. G., 1973, "Thin Film Lubrication Theory for Newtonian Fluids with Surfaces Possessing Striated Roughness or Grooving," ASME *Journal of Lubrication Technology*, Oct., pp. 484–489.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Grimm, R. J., 1976, "Squeezing Flows of Newtonian Liquid Films, An Analysis Including Fluid Inertia," *Appl. Sci. Res.*, Vol. 32, pp. 149–166.

Hasegawa, E., and Izuchi, H., 1984, "On Steady Flow Through A Channel Consisting of An Uneven and A Plane Wall," *Bulletin of the JSME*, Vol. 27, pp. 1631–1636.

Kuhn, K. C., and Yates, C. C., 1964, "Fluid Inertia Effects on the Film Pressure Between Axially Oscillating Parallel Circular Plates," *Trans. ASME*, Vol. 7, pp. 294–303.

Michell, A. G. M., 1950, "Lubrication—Its Principles and Practice," Blackie and Son, London, pp. 281–291.

Pereyra, V., 1978, "PASVA3: An Adaptive Finite Difference Fortran Program for First Order Nonlinear Ordinary Boundary Problems," Lect. Notes Comp. SC 76: 67–88, Springer-Verlag, Berlin.

Sun, D., and Cheng, K. K., 1977, "First Effects of Stokes Roughness on Hydrodynamic Lubrication," ASME *Journal of Lubrication Technology*, Jan., pp. 2–9.

Tichy, J. A., and Modest, M. F., 1978, "Squeeze Film Flow Between Arbitrary Two Dimensional Surfaces Subject to Normal Oscillations," ASME *Journal of Lubrication Technology*, July, Vol. 100, p. 136.

Tichy, J. A., and Winer, W. O., 1970, "Inertial Considerations in Parallel Circular Squeeze Film Bearings," ASME *Journal of Lubrication Technology*, Vol. 92, No. 4, ASME, New York, pp. 588–592.

Tonder, K., 1977, "Mathematical Verification of the Applicability of Modified Reynolds Equations to Striated Roughness Surfaces," *Wear*, Vol. 44, pp. 329–343.

Wang, C. Y., 1976, "Parallel Flow Between Corrugated Plates," *Journal of Engineering Mechanics Division*, Vol. 201, No. EM6, ASCE, New York, pp. 1088–1090.

Wang, C. Y., 1978, "Drag Due To A Striated In Boundary Slow Couette Flow," *Physics of Fluids*, Vol. 21, pp. 697–698.

White, J. W., 1980, "Surface Roughness Effects on the Load Carrying Capacity of Very Thin Compressible Lubricating Films," ASME *Journal of Lubrication Technology*, Vol. 102, Oct., pp. 445–451.

APPENDIX

Numerical Solution of the Differential Equations

To solve the complicated differential equations of this work, e.g., equations (26) and (27), the DVCPR numerical routine was used. This routine solves a system of ordinary differential equations with boundary conditions at two points, using a variable step-size finite difference method with deferred corrections. Therefore, equations (26) and (27) are first rewritten as a set of ordinary differential equations as the numerical scheme requires:

$$Y_1 = \bar{\psi}'^r$$

$$Y_2 = F_1 = \bar{\psi}'^r_\eta$$

$$Y_3 = F_2 = \bar{\psi}'^r_{\eta\eta}$$

$$Y_4 = F_3 = \bar{\psi}'^r_{\eta\eta\eta}$$

$$F_4 = \mathrm{Re}^*(-2\pi\alpha(\eta)Y_7 + 2\pi\beta Y_5) = \bar{\psi}'^r_{\eta\eta\eta\eta}$$

$$Y_5 = \bar{\psi}'^i$$

$$Y_6 = F_5 = \bar{\psi}'^i_\eta$$

$$Y_7 = F_6 = \bar{\psi}'^i_{\eta\eta}$$

$$Y_8 = F_7 = \bar{\psi}'^i_{\eta\eta\eta}$$

$$F_8 = \mathrm{Re}^*(2\pi\alpha(\eta)Y_3 - 2\pi\beta Y_1 - 2\pi\gamma(\eta)) = \bar{\psi}'^i_{\eta\eta\eta\eta}.$$

The basic discretization used in this routine is the trapezoidal rule over a possibly nonuniform mesh. This mesh is chosen adaptively in order to make the local error approximately the same size everywhere. Higher-order discretizations are obtained by deferred corrections. The resulting algebraic system is solved by Newton's method with a step control, while the linearized sparse system is solved by a special form of Gaussian elimination that preserves the sparseness (Pereyra (1978)).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Stability of Liquid Film Flow Down an Oscillating Wall

Ronald J. Bauer
Vitro Corp.,
Silver Spring, MD 20906-2972

C. H. von Kerczek
Mechanical Engineering Department,
University of Maryland Baltimore County,
Baltimore, MD 21228
Mem. ASME

The stability of a liquid film flowing down an inclined oscillating wall is analyzed. First, the linear theory growth rates of disturbances are calculated to second order in a disturbance wave number. It is shown that this growth rate is simply the sum of the same growth rate expansions for a nonoscillating film on an inclined plate and an oscillating film on a horizontal plate. These growth rates were originally calculated by Yih (1963, 1968). The growth rate formula derived here shows that long wavelength disturbances to a vertical falling film, which are unstable at all nonzero values of the Reynolds number when the wall is stationary, can be stabilized by sufficiently large values of wall oscillation in certain frequency ranges. Second, the full time-dependent stability equations are solved in terms of a wall oscillation amplitude expansion carried to about 20 terms. This expansion shows that for values of mean flow Reynolds number less than about ten, the wall oscillations completely stabilize the film against all the unstable disturbances of the steady film.

Introduction

It is well known that the flow of a liquid film down a vertical wall is unstable (Yih, 1963). The instability is manifested as temporally growing free surface waves whose phase velocity is equal to the mean flow velocity at the free surface. It is also well known that a vertically standing pendulum is unstable but that it can be stabilized by suitably oscillating the pendulum in the vertical direction. This fact suggests that the liquid film might also be stabilized by oscillating the wall in the direction of the mean flow. This possibility is examined in this study of solving the linearized unsteady film flow stability problem.

The stability of a liquid film on a horizontal oscillating flat plate was examined by Yih (1968). He showed that for small values of the wave number of disturbances, this flow could be unstable at certain values of the imposed wall oscillation frequency and amplitude. Von Kerczek and Davis (1974) showed that the Stokes oscillating boundary layer itself is probably stable up to very large values of Reynolds number. Thus, instabilities of oscillating films are due to temporally growing free surface waves and are not shear instabilities of the underlying Stokes layer.

This study solves approximately the linearized stability problem for an inclined film on an oscillating wall in two ways. The first way is by employing Yih's (1963, 1968) small wave number approximation. This method yields a formula for the growth rate of disturbances of long wavelengths. This formula is accurate to second order in the wave number for all values of wall oscillation amplitude and frequency and finite values of the product of the wave number and Reynolds number. Most interesting is that the growth rate formula derived here is simply the sum of the growth rate formula of Yih's (1963) steady falling film stability problem and Yih's (1968) horizontal oscillating film stability problem.

The second way that the stability problem is solved here is by a high-order series expansion in the wall oscillation amplitude. This series is summed with the aid of Shanks' transformation (see Shanks, 1955; or Van Dyke, 1984) to yield accurate values of disturbance growth rates in a range of moderate (but significant) values of the wave number and Reynolds number.

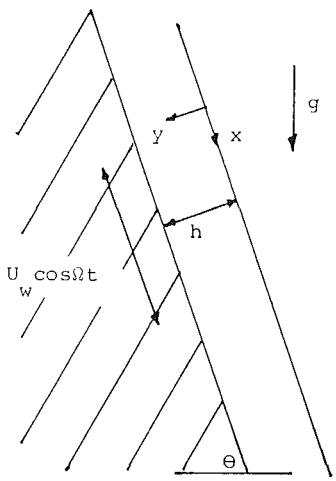


Fig. 1 Schematic diagram of the film flow down an inclined oscillating wall

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

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technical Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 13, 1988; final revision, Dec. 4, 1989.

Copyright © 1991 by ASME

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The result of both methods of analysis is to show that a vertical film can indeed be stabilized by vertical wall oscillations in a certain frequency range and a sufficiently large amplitude.

Formulation

A schematic of the flow configuration is shown in Fig. 1. The problem is formulated in terms of the film thickness h as the length scale, the film speed U at the free surface as the velocity scale, h/U as the time scale, and ρU^2 as the stress scale. Henceforth it is assumed that all variables have been rendered dimensionless by suitable combinations of the above scales. The following parameters emerge from this scaling: $U=h^2g\ \sin(\theta)/2\nu$; the Reynolds number $R=Uh/\nu$; the reciprocal Weber number $S=(T/\rho)\,(2/g\nu^4\ \sin(\theta))^{1/3}$ where T is the surface tension; the dimensionless frequency $\sigma=\Omega h/U$; the dimensionless wall oscillation amplitude $\epsilon=U_w/U$; and the frequency parameter (the Womersly number) $\beta=\sqrt{\sigma R/2}$, which represents the ratio of the layer thickness h to Stokes layer thickness $\delta=\sqrt{2\nu/\Omega}$. For the purpose of illustration it is assumed (as by Yih) that the liquid is distilled water at standard conditions for which $S=4900/R^{5/3}$.

The flow is assumed to be two-dimensional in the x, y coordinate system shown in Fig. 1. The velocity components in the x, y directions and the free surface deflection are generally denoted by u, v, ζ, respectively, with various subscripts to denote the special cases. In particular, the subscript b denotes the basic state, the subscript s denotes a steady flow quantity and the subscript w denotes a purely oscillatory quantity associated with the wall oscillation. The basic flow down the oscillating wall is described by

$$\mathbf{V}=(u_b(y,\ t),\ 0,\ 0),\quad P=gh\ \cos(\theta)/U^2,\quad \zeta_b=0 \tag{1}$$

where

$$u_s(y)=1-y^2 \tag{2a}$$

$$u_w(y,t)=\Re\{(\cosh[(1+i)\beta y]/\cosh[(1+i)\beta])\exp(i\sigma t)\}. \tag{2b}$$

The linear stability of two-dimensional disturbances of wave number α can be described in terms of a complex stream function ϕ and a complex deflection η in such a way that the disturbance velocities u, v and free surface elevation ζ are given by the equations

$$u=-\Re\{(\partial\phi/\partial y)\ \exp\ (i\alpha x)\} \tag{3a}$$

$$v=\Re\{i\alpha\phi\ \exp\ (i\alpha x)\} \tag{3b}$$

$$\zeta=\Re\{\eta\ \exp\ (i\alpha x)\}. \tag{3c}$$

By linearizing the Navier-Stokes equations and the governing boundary conditions in the usual way of linear stability theory (Yih 1963, 1968 or Bauer, 1986), the unsteady Orr Sommerfeld equation

$$L\partial\phi/\partial t+i\alpha u_bL\phi-i\alpha\phi D^2u_b=(1/R)L^2\phi \tag{4a}$$

where $D=d/dy$ and $L=D^2-\alpha^2$ with boundary conditions

$$\phi=\partial\phi/\partial y=0\quad \text{at } y=1, \tag{4b}$$

$$\partial\eta/\partial t+i\alpha(u_b\eta+\phi)=0\quad \text{at}\quad y=0 \tag{4c}$$

$$\partial^2\phi/\partial y^2+\eta\partial^2u_b/\partial y^2+\alpha^2\phi=0\quad \text{at}\quad y=0 \tag{4d}$$

$$\partial^2\phi/\partial t\partial y+i\alpha u_b\partial\phi/\partial y-i\alpha\left(\alpha^2S+\frac{2}{R}\cot\theta\right)\eta=$$
$$(1/R)\,(D^3-3\alpha^2D)\phi\quad \text{at}\quad y=0 \tag{4e}$$

must be solved. The boundary condition (4c) is the kinematic free surface condition and 4(d), 4(e) are the dynamic (stress continuity) free surface conditions.

Small Wave Number Expansion

The small wave number analysis of Yih (1963, 1968) is used here to establish an approximate analytical estimate of the stability characteristics of the oscillating film. Bauer (1968) showed by direct solution of the small wave number approximation of equations (4) that the growth rate of disturbances to the oscillatory falling film is merely the sum of the growth rate of disturbances in the steady falling film as derived by Yih (1963) and the growth rate of the horizontal oscillating film as derived by Yih (1968). This derivation requires a formidable amount of algebra. In the following analysis it is shown directly that the growth rate of a disturbance in the oscillating film must be the sum of the growth rates derived by Yih.

The time periodicity of the coefficients of equations (4), according to Floquet theory (Coddington and Levinson, 1963), indicates that the solution (ϕ, η) can be expressed in the form

$$\phi=\psi(y,t)\exp(\lambda t)\quad \text{and}\quad \eta=\eta\ \exp(\lambda t) \tag{5}$$

where (ψ, μ) are $2\pi/\sigma$ periodic functions of time. The Floquet exponent λ determines the stability ($\Re\lambda<0$) or instability ($\Re\lambda>0$) of the basic state.

The three unknown quantities (ψ, μ, λ) are expanded for small values of wave number α as follows:

$$\psi=\psi_0+\alpha\psi_1+\alpha^2\psi_2+\ldots \tag{6a}$$

$$\mu=\mu_0+\alpha\mu_1+\alpha^2\mu_2+\ldots \tag{6b}$$

$$\lambda=\lambda_0+\alpha\lambda_1+\alpha^2\lambda_2+\ldots \tag{6c}$$

Equations (5) and (6) are then substituted into equations (4). The usual sorting of terms according to powers of α yields the various problems that need to be solved. For later convenience the following generic operators are defined:

$$\mathcal{L}(\psi,\ \lambda)=\lambda D^2\psi+dD^2\psi/dt-(1/R)D^4\psi \tag{7a}$$

$$B_1(\lambda,\ \mu)=d\mu/dt+\lambda\mu \tag{7b}$$

$$B_2(\psi,\ \mu)=D^2\psi+\mu D^2u_b \tag{7c}$$

$$B_3(\lambda,\ \psi)=dD\psi/dt+\lambda D\psi-(1/R)D^3\psi. \tag{7d}$$

Thus, small wave number expansion terms are:

$O(1)$:
$$\mathcal{L}(\psi_0,\ \lambda_0)=0 \tag{8a}$$

with boundary conditions

$$\psi_0=D\psi_0=0\quad \text{at}\quad y=1 \tag{8b}$$

$$B_1(\lambda_0,\ \mu_0)=0\quad \text{at}\quad y=0 \tag{8c}$$

$$B_2(\psi_0,\ \mu_0)=0\quad \text{at}\quad y=0 \tag{8d}$$

$$B_3(\lambda_0,\ \psi_0)=0\quad \text{at}\quad y=0. \tag{8e}$$

$O(\alpha)$:
$$\mathcal{L}(\psi_1,\ \lambda_0)=-\lambda_1D^2\psi_0-i(u_bD^2\psi_0-\psi_0D^2u_b) \tag{9a}$$

with boundary conditions

$$\psi_1=D\psi_1=0\quad \text{at}\quad y=0 \tag{9b}$$

$$B_1(\lambda_0,\ \mu_1)=-\lambda_1\mu_0-iu_b\mu_0-i\psi_0\quad \text{at}\quad y=0 \tag{9c}$$

$$B_2(\psi_1,\ \mu_1)=0\quad \text{at}\quad y=0 \tag{9d}$$

$$B_3(\lambda_0,\ \psi_1)=-\lambda_1D\psi_0$$
$$-i\left[u_bD\psi_0-\left(\alpha^2S+\frac{2}{R}\cot\theta\right)\mu_0\right]\quad \text{at}\quad y=0. \tag{9e}$$

At order α^2 the problem is similar to one at order α, and in order to be brief it need not be stated here. It suffices to note that the kinematic free surface condition (4c) to order α^2 is

$$B_1(\lambda_0,\ \mu_2)=-\lambda_2\mu_0-\lambda_1\mu_1-iu_b\mu_1-i\psi_1. \tag{10}$$

This boundary condition is sufficient to determine the value of λ_2, and it only requires the knowledge of the solutions at $O(1)$ and $O(\alpha)$.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The values of λ_i, for $i=1, 2, \ldots$, are determined from equations (8*c*), (9*c*), and (10) by applying the periodicity condition

$$\int_0^{2\pi/\sigma} B_1(\lambda_0, \mu_i)dt = 0. \tag{11}$$

It is now possible to proceed with the analysis to determine the terms λ_i. By applying equation (11) to boundary condition (8*c*), it is evident that μ_0 is a constant and that $\lambda_0 = 0$. For the sake of definiteness it is permissible to normalize so that $\mu_0 = 1$. It is then an easy matter to calculate ψ_0, which turns out to be

$$\psi_0 = (1-y)^2 + \frac{\epsilon}{2}\left[\frac{1-\cosh\gamma(1-y)}{\cosh\gamma}\right]\exp(i\sigma t) + cc \tag{12}$$

where $\gamma = (1+i)\beta$ and *cc* stands for the complex conjugate of the preceding term.

The term λ_1 is now calculated by using equations (9*c*) and (12) in the periodicity condition (11). The result of this calculation is

$$\lambda_1 = -2i. \tag{13}$$

The solution of the equation (9*c*) is

$$\mu_1 = -\frac{\epsilon}{2\sigma\cosh\gamma}e^{i\sigma t} - cc. \tag{14}$$

In order to calculate λ_2 the solution ψ_1 must be computed. A large amount of algebra can be avoided by exploiting the linearity of equation (15) and that u_b and ψ_0 can be split into a steady and a purely sinusoidal part as follows:

$$u_b = u_s + u_w e^{i\sigma t} + cc \tag{15a}$$

$$\psi_0 = \psi_{0s}(y) + \psi_{0w}(y)e^{i\sigma t} + cc. \tag{15b}$$

By applying the periodicity condition (11) to equation (10), one obtains

$$\lambda_2 = -i(u_b|_{y=0}\mu_1 + \psi_1|_{y=0})_s \tag{16}$$

where the subscript *s* denotes that only the steady part of the quantity needs to be considered. It can be seen from equation (16) that only the part ψ_{1s} of the solution ψ_1 and the steady part of the product $u_b\mu_1$ need to be computed.

Equation (9*a*) reduces for ψ_{1s} to

$$D^4\psi_{1s} = -\lambda_1 D^2\psi_{0s} - i(u_b D^2\psi_0 - \psi_0 D^2 u_b)_s \tag{17a}$$

with boundary conditions

$$\psi_{1s} = D\psi_{1s} = 0 \quad \text{at} \quad y=1 \tag{17b}$$

$$D^2\psi_{1s} = -(\mu_1 D^2 u_b)_s \quad \text{at} \quad y=0 \tag{17c}$$

$$D^3\psi_{1s} = \lambda_1 R D\psi_{0s} - iR\Lambda + iR(u_b D\psi_0)_s \quad \text{at} \quad y=0 \tag{17d}$$

where $\Lambda = \alpha^2 S + (2/R)\cot\theta$.

The steady parts of the product terms in equations (17) are

$$(u_b D^2\psi_0 - \psi_0 D^2 u_b)_s = (u_s D^2\psi_{0s}^* + cc - \psi_{0s}^* D^2 u_s - cc) + (u_w D^2\psi_{0w}^* + cc - \psi_{0w}^* D^2 u_w - cc) \tag{18a}$$

$$(D^2 u_b \mu_1)_s = -D^2 u_w \mu_1 + cc \quad \text{at} \quad y=0 \tag{18b}$$

$$(u_b D\psi_0)_s = u_s D\psi_{0s}^* + cc + u_w D\psi_{0w}^* + cc \quad \text{at} \quad y=0 \tag{18c}$$

where the asterisk superscript denotes complex conjugation. Equations (18) indicate clearly that the inhomogeneous terms of equations (17) separate into additive terms that are individually the steady and purely oscillatory terms of the $O(1)$ problem, equations (8). Furthermore, by exploiting the linearity of equations (17) the solution ψ_{1s} also can be divided into two parts

$$\psi_{1s} = \psi_{1ss} + \psi_{1sw} \tag{19}$$

where ψ_{1ss} depends only on the steady functions u_s and ψ_{0s}, and ψ_{1sw} depends only on the oscillatory functions u_w and ψ_{0w}. Then ψ_{1ss} satisfies

$$D^4\psi_{1ss} = -\lambda_1 D^2\psi_{0s} - i(u_s D^2\psi_{0s}^* + cc - \psi_{0s}^* D^2 u_s - cc) \tag{20a}$$

with boundary conditions

$$\psi_{1ss} = D\psi_{1ss} = 0 \quad \text{at} \quad y=1 \tag{20b}$$

$$D^2\psi_{1ss} = 0 \quad \text{at} \quad y=0 \tag{20c}$$

$$D^3\psi_{1ss} = \lambda_1 R D\psi_{0s}^* - iR\Lambda + iR(u_s D\psi_{0w}^* + cc) \quad \text{at} \quad y=0 \tag{20d}$$

and ψ_{1sw} satisfies the equations

$$D^4\psi_{1sw} = -i(u_w D^2\psi_{0w}^* + cc - \psi_{0w}^* D^2 u_w - cc) \tag{21a}$$

with boundary conditions

$$\psi_{1sw} = D\psi_{1sw} = 0 \quad \text{at} \quad y=1 \tag{21b}$$

$$D^2\psi_{1sw} = -\mu_1 D^2 u_w \quad \text{at} \quad y=0 \tag{21c}$$

$$D^3\psi_{1sw} = iR(u_w D\psi_{0w}^* + cc) \quad \text{at} \quad y=0. \tag{21d}$$

Furthermore, the exponent λ_2 can be split into two parts as

$$\lambda_2 = \lambda_{2s} + \lambda_{2w} \tag{22}$$

where

$$\lambda_{2s} = -i\psi_{1ss} \quad \text{at} \quad y=0 \tag{23}$$

and

$$\lambda_{2w} = -i(u_s\mu_1 + \psi_{1sw}) \quad \text{at} \quad y=0. \tag{24}$$

Equations (20) and (23) can be recognized to be Yih's (1963) equations for the growth rate of long-wave disturbances on a steady inclined film (albeit in different notation). Yih's result in our notation gives

$$\lambda_{2s} = \frac{8R}{15} - R\left(\frac{1}{3}\alpha^2 S + \frac{2}{3R}\cot\theta\right). \tag{25a}$$

It is somewhat more difficult to relate equations (21) and (24) to Yih's (1968) horizontal oscillating layer results because he uses a completely different scaling. Furthermore, when considering the horizontal layer in isolation, the stability formulation contains a constant term equivalent to (but not the same as) our term Λ in equations (17*d*) and (20*d*). Yih denotes this term as the square of an inverse Froude number. The way in which the terms proportional to the steady and oscillatory parts of the basic state have been separated in the above formulation requires that Yih's (1968) term, proportional to the square of inverse Froude number, be dropped in the comparison of his equations (ic, iic, iiic, and ivc following equation 56 on p. 747) to our equations (21) and (24). A detailed comparison of Yih's formulation (with the appropriate reconciliation of notation) then yields for the present problem that

$$\lambda_{2w} = -\frac{3R\epsilon^2}{16\beta^2(\sinh^2\beta + \cos^2\beta)^2}[\beta\cosh(2\beta)\sin(2\beta) + \beta\sinh(2\beta)\cos(2\beta) - \sinh(2\beta)\sin(2\beta)]. \tag{25b}$$

The discussion of the growth rate formula (25) will be postponed until after the full time-dependent stability problem (4) is solved by an expansion in oscillation amplitude ϵ.

Expansion in Oscillation Amplitude

Equations (4) will now be solved by the oscillation amplitude expansion developed by von Kerczek (1982). In this method the stream function ϕ is first transformed to the interval $z\epsilon[-1, 1]$ by the equation $z = 2y - 1$. The stream function ϕ is expanded in a Chebyshev series whose coefficients f_k are functions of time t. Thus,

$$\phi(y,t) = \sum_{k=0}^{K-1} f_k(t)T_k(z). \tag{26}$$

This expansion, together with the boundary conditions (4*b*) – (4*e*), converts the Orr Sommerfeld stability problem (4)

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Table 1 Growth rate coefficients

$\alpha = 0.08$ $R = 10$ $\beta = 3$

k	real(λ)	imag(λ)
0	.13723D-01	−.14977D+00
2	−.50205D-02	−.13578D−02
4	.27119D-03	.17374D−03
6	−.13402D-05	−.13319D−05
8	.19132D-07	−.34748D−06
10	.49822D-08	.36853D−08
12	−.69562D-09	.28261D−09
14	.50372D-11	−.21826D−10
16	.15581D-11	.14284D−11
18	−.81720D-13	.17751D−13
20	.20295D-14	−.65403D+14

into a system of ordinary differential equations in time t of the form

$$d[A(t)\mathbf{F}]/dt = B(t)\mathbf{F} \tag{27}$$

where $A(t)$ and $B(t)$ are $K \times K$ matrices and $\mathbf{F} = (\mathbf{f}_0, f_1, \ldots, f_{K-1})^T$ is the K-dimensional column vector of the Chebyshev coefficients. The details of the conversion of equations (4) to the equations (27) can be found in Bauer (1986).

The matrices A and B are $2\pi/\sigma$ time periodic, and by the Floquet theorem the solution $\mathbf{F}$ has the form

$$\mathbf{F} = \mathbf{P}(t)\exp(\lambda t) \tag{28}$$

where $\mathbf{P(t)}$ is a $2\pi/\sigma$ time-periodic vector function. The oscillatory flow is thus stable or unstable according to $\Re\lambda < 0$ or > 0, respectively.

The matrices A and B in equation (27) have the form

$$A = I + \epsilon[A^+\exp(i\sigma t) + A^-\exp(-i\sigma t)] \tag{29a}$$

$$B = B_0 + \epsilon[B^+\exp(i\sigma t) + B^-\exp(-i\sigma t)] + \epsilon^2[B^{++}\exp(2i\sigma t) + B_{00} + B^{--}\exp(-2i\sigma t)]. \tag{29b}$$

The matrices B_0, A^+, A^-, etc., are complex constant matrices. By substituting the expansions

$$\mathbf{P} = \mathbf{P}_0 + \mathbf{P}_1\epsilon + \mathbf{P}_2\epsilon^2 + \ldots \tag{30a}$$

$$\lambda = \lambda_0 + \lambda_1\epsilon + \lambda_2\epsilon^2 + \ldots, \tag{30b}$$

and the matrices (29*a*), (29*b*) into equations (28) and (27), respectively, and then collecting terms in powers of ϵ, there results a sequence of constant coefficient differential equations of the form

$$d\mathbf{P}_n/dt - B_0\mathbf{P}_n = \Psi\{\mathbf{P}_k\exp(\pm ik\sigma t),\ \lambda_k; \exists k < n; \lambda_n\} \tag{31}$$

for $n = 1, 2, \ldots$. The right side of equation (31) denotes a vector function of all the lower-order terms $\mathbf{P}_k$, λ_k, the Fourier components $\exp(\pm ik\sigma t)$ and the current Floquet exponent coefficient λ_n. The details of this function Ψ can be found in Bauer (1986).

The requirement that each of the solutions $\mathbf{P}_n$ be $2\pi/\sigma$ periodic forces the function Ψ to be likewise periodic at each order. This allows the sequential determination of the coefficients λ_n. It should be noted that all the odd-numbered coefficients λ_n are zero but not the odd-numbered vectors $\mathbf{P}_n$ (see Bauer, 1986). Hence, the expansion for λ will only contain even powers of ϵ.

Up to 21 ($N = 20$) terms of the series (31) can be computed (on our modest PC). The resulting 11-term series for λ are then summed with the aid of the following Shanks' transformation (see Van Dyke, 1984):

$$S_{m,0} = \sum_{k=0}^{m} \lambda_{2k}\epsilon^{2k} \tag{32a}$$

$$S_{m,j} = \frac{(S^2_{m,j-1} - S_{m+1,j-1}S_{m-1,j-1})}{(2S_{m,j-1} - S_{m+1,j-1}S_{m-1,j-1})}. \tag{32b}$$

The accepted value of $\lambda = S_{M,M}$ for $M = N/4$; the value of S at the apex of the triangular array (32). This value of λ is deemed to be reliable to at least two significant figures if the entries on the main descending diagonal of the array $S_{m,j}$ have converged (in the Cauchy sense) to two significant figures.

An example series expansion of λ is given in Table 1 for $R = 10$, $\alpha = 0.8$, $\beta = 3$, and $N = 20$. In Table 1 the coefficients have been scaled to avoid computer underflow so that the series with the listed coefficients proceeds in powers of $(\epsilon/10)^k$.

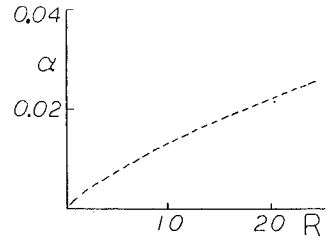


Fig. 2 The neutral stability curve for steady flow of a film on a vertical wall

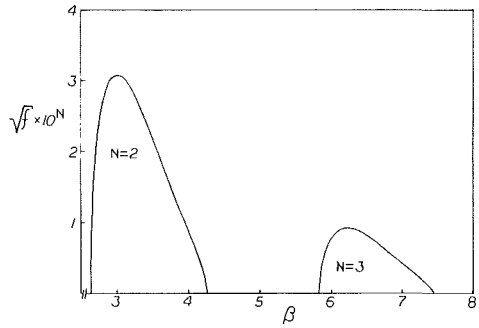


Fig. 3 The oscillatory stability function $\sqrt{f(\beta)}$, in the range where $f \geq 0$, for determining the effects of wall oscillation on disturbance growth rates (see equation (33))

Discussion of the Results

For the sake of definiteness and in order to conform to Yih's and others' results we will apply our results to the case of distilled water at standard conditions for which $S = 4900/R^{5/3}$.

It is useful to first consider the stability characteristics of a steady film flowing down an inclined plane. Although the above developments are applicable to films of any inclination, only vertical films are considered for the purpose of illustration. Figure 2 shows the neutral stability curve for the steady vertical water film. For values of α to the left of the curve, the film in unstable. It should be noted that the R-axis is part of the neutral curve. At small values of Reynolds number only very long wavelength (i.e., very low frequency) disturbances are unstable.

The growth rate equation (25) will be discussed first. It is most informative to write this equation in the form

$$\lambda_2/R = K(R,\alpha,\theta) - \epsilon^2 f(\beta) \tag{33}$$

where K and f are clearly defined by comparing this equation with equations (25). The neutral curve of ϵ versus β for fixed values of R, α, and θ is obtained by setting $\lambda_2 = 0$ in equation (33). This yields the equation

$$\epsilon = \sqrt{[K(R,\alpha,\theta)/f(\beta)]}. \tag{34}$$

This equation is valid as long as the product αR is sufficiently small. Also, only positive values of f are significant. It is worth noting that negative values of f serve to destabilize the film and are the relevant values of f in considering a horizontal

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

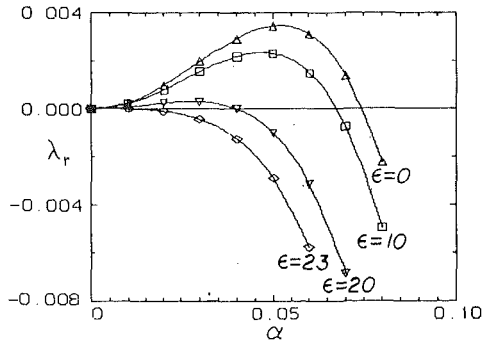


Fig. 4 Growth rate versus wave number for $R=5$, $\beta=3$ and various values of ϵ

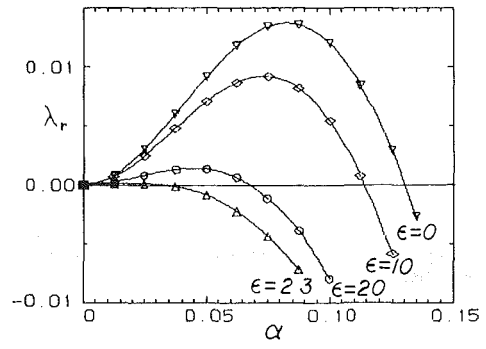


Fig. 5 Growth rate versus wave number for $R=10$, $\beta=3$ and various values of ϵ

liquid layer undergoing horizontal oscillations (see Yih's 1968 Fig. 2 on p. 749). Thus, in considering a nonvertical steady film, which has a critical value of Reynolds number given by $R=(5/4)\cot\theta$ at $\alpha=0$, the oscillations will decrease this critical value for the oscillation frequency range in which f is negative. The more interesting situation here is when f is positive and the imposed oscillations tend to stabilize the film.

The function $\sqrt{f}$ is plotted in Fig. 3 for positive values of f versus β. Thus, given any film inclination angle θ, formula (34) and Fig. 3 can be used to compute the required value of oscillation amplitude ϵ that stabilizes the film to long-wave disturbances. It is to be noted that the optimal oscillation frequency for stabilization corresponds to $\beta\sim 3$. The vertical film ($\theta=\pi$) requires an amplitude $\epsilon=23.7$ for stabilization of long-wave disturbances, $\alpha\to 0$. This large value of oscillation amplitude is required because of the mismatch of oscillation frequency and disturbance frequency which is approaching zero with wave number. However, the growth rate of the long-wave disturbances also approaches zero with wave number. Thus, these disturbances remain neutrally stable. (Recall that the actual growth rate of a disturbance is $\lambda=\alpha^2\lambda_2$.) What changes is the sign (from positive to negative) of the slope of the growth rate with respect to α^2 holding R constant. It is interesting to note that the optimal long-wave stabilization frequency of $\beta=3$ and amplitude $\epsilon=23.7$ holds for all vertical films. For an actual example, a real film of pure water flowing at Reynolds number $R=10$, the film thickness is 0.00475 in. (0.0121 cm) and thus requires an oscillation amplitude of about 0.25 in. (0.635 cm) and oscillation frequency of 1.3 hz for complete stabilization to long wavelength disturbances.

We remark that there are an infinite number of intervals on the β-axis on which $f(\beta)$ is positive, and thus the film can be stabilized. However, as shown in Fig. 3 for the second such interval, the magnitude of f is so small that enormous values of oscillation amplitude are required. For example, at $\beta=6.2$ the amplitude has to be $\epsilon\sim 800$. Such oscillations are clearly impractical.

The complete time-dependent stability equations (4) were solved by the amplitude expansion method described in a previous section. A sample of the expansion results was given there. Figures 4 and 5 show typical results for the vertical water film of the modification of the steady flow instability growth rates as a function of the wave number α for various values of wall oscillation amplitude ϵ. These results are for $R=5$ and 10 and $\beta=3$ in Figs. 4 and 5, respectively. It can be seen that at each of the values of Reynolds number, the flow becomes increasingly more stable with increases of the value of wall oscillation amplitude ϵ in the sense that the interval in α for which instability occurs is narrowed, and the maximum value of the growth rate is lowered. Furthermore, it is evident that the long wavelength (small α) disturbances are the most difficult to stabilize. However, with a wall oscillation amplitude of $\epsilon>23.7$ at wall frequency of $\beta=3$, the film becomes completely stabilized at least for $R\le 10$. Calculations were made at $R=15$, but 21 terms of the series were insufficient to yield converged values (to at least three figures) of the growth rate expansion, even using the Shanks transformation. These results did converge for values of $\alpha<0.04$ and did indicate that a wall oscillation amplitude of at least 23.7 was required for stabilization.

What is most interesting here is that the small wave number asymptotic result (25) seems to be a sufficient condition for complete stabilization of the vertical water film. Although one cannot be definitive about this statement, until it is checked at larger values of Reynolds number, it can be useful as a guide or preliminary estimate of what might be required to dynamically stabilize such films. It is unfortunate that the amplitude expansion method can only be used at fairly small values of Reynolds number unless a very large number of terms is calculated. Thus, in order to examine the effects of wall oscillation in the Reynolds number range of maximum growth rate (about 35 for the vertical water film), the numerical time integration method of von Kerczek and Davis (1974) should be used. Such an endeavor is underway.

References

Bauer, R. J., 1986, "Effect of Oscillatory Excitation on Film Flow Stability," Ph.D. Thesis, The Catholic University of America, (#8613469 University Microfilms International), Ann Arbor, Mich.

Coddington, E. A., and Levinson, N., 1955, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

Shanks, D., 1955, "Non-linear Transformations of Divergent and Slowly Convergent Sequences," *J. Math. and Phys.*, Vol. 34, pp. 1–42.

Van Dyke, M., 1984, "Computer-Extended Series," *Annual Review of Fluid Mechanics*, Vol. 16, pp. 287–309.

von Kerczek, C. H., and Davis, S. H., 1974, "Linear Stability Theory of Oscillatory Stokes Layers," *J. Fluid Mech.*, Vol. 62, pp. 753–773.

von Kerczek, C. H., 1982, "The Instability of Oscillatory Plane Poiseuille Flow," *J. Fluid Mech.*, Vol. 116, pp. 91–114.

Yih, C. S., 1963, "Stability of Liquid Flow Down an Inclined Plane," *Physics of Fluids*, Vol. 6, pp. 321–334.

Yih, C. S., 1968, "Instability of a Horizontal Layer on an Oscillating Plane," *J. Fluid Mech.*, Vol. 31, pp. 737–751.

Downloaded 03 May 2010 to 171.66.16.248. 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.

Universal Relations for Elastically Isotropic Elastic-Plastic Materials

M. B. Rubin[1,2] and R. Chen[1]

1 Introduction

In the theory of nonlinear elasticity it is well known that it is possible to develop universal relations which, in general, connect kinematic and kinetic quantities and which are valid for any isotropic elastic material. One of the most common examples of a universal relation is the one developed for simple shear (see Truesdell and Noll, 1965, Sec. 54). Such relations can be used in connection with experiments to either check the accuracy of the experimental methods or to check the validity of the assumption of isotropy. These universal relations can also be used in computations or analytical work to provide simpler expressions for certain restrictions.

By way of background let: $\mathbf{X}$ denote the position of a material point in the reference configuration; $\mathbf{x}$ denote the position of the same material in the present configuration at time t; and let the deformation tensors $\mathbf{F}$, $\mathbf{C}$, $\mathbf{B}$, and the dilatation I_3 be defined by

$$\mathbf{F} = \partial \mathbf{x}/\partial \mathbf{X}, \tag{1a}$$

$$\mathbf{C} = \mathbf{F}^T\mathbf{F},\ \mathbf{B} = \mathbf{F}\mathbf{F}^T,\ I_3 = \det \mathbf{C}. \tag{1b,c,d}$$

Furthermore, let $\mathbf{T}$ be the Cauchy stress tensor; $\mathbf{T}'$ be its deviatoric part, p be the pressure, and $\mathbf{S}$ be the symmetric Piola Kirchhoff stress tensor, such that

$$\mathbf{T} = -p\mathbf{I} + \mathbf{T}',\quad \mathbf{T}'\cdot\mathbf{I} = 0 \tag{2a,b}$$

$$\mathbf{S} = -pI_3^{1/2}\mathbf{C}^{-1} + \mathbf{S}',\ \mathbf{S}'\cdot\mathbf{C} = 0, \tag{2c,d}$$

$$p = -(1/3)\mathbf{T}\cdot\mathbf{I} = -(1/3)I_3^{-1/2}\mathbf{S}\cdot\mathbf{C}, \tag{2e}$$

$$\mathbf{T} = I_3^{-1/2}\mathbf{F}\mathbf{S}\mathbf{F}^T,\ \mathbf{T}' = I_3^{-1/2}\mathbf{F}\mathbf{S}'\mathbf{F}^T, \tag{2f,g}$$

where $\mathbf{I}$ is the unity tensor and $\mathbf{A}\cdot\mathbf{B} = \mathrm{tr}(\mathbf{A}\mathbf{B}^T)$ is the inner product between two tensors.

Recently, Beatty (1987) analyzed a class of universal relations for isotropic elastic materials which is characterized by the simple tensor equation

$$\mathbf{B}\mathbf{T} = \mathbf{T}\mathbf{B}. \tag{3}$$

Beatty (1987) showed that the three independent scalar equations associated with equation (3) generate many of the known universal relations as well as some new ones. Alternatively, equation (3) may be rewritten in terms of $\mathbf{C}$ and $\mathbf{S}$ by using equations (1) and (2*f*) to obtain

$$\mathbf{C}\mathbf{S} = \mathbf{S}\mathbf{C}. \tag{4}$$

It also follows from equations (2*c*) and (4) that

$$\mathbf{C}\mathbf{S}' = \mathbf{S}'\mathbf{C}. \tag{5}$$

Independently, Rubin (1988a) used equation (5) to develop a universal relationship for pure shear ($T_{12} = T_{21} \neq 0$, all other $T_{ij} = 0$). However, Rubin (1988a) was unaware of the general nature of equation (5) until reading Beatty's paper (1987).

The objective of this note is to consider a class of universal relations which are valid for any elastically isotropic elastic-plastic or elastic-viscoplastic materials in a rather general class of hyperelastic-type constitutive equations for stress.

2 Constitutive Equations

Using the thermodynamic procedure proposed by Green and Naghdi (1977, 1978) it can be shown (Rubin, 1986) that a rather general class of elastic-viscoplastic materials is described by the independent quantities

$$U_1 = \{\mathbf{C},\ \mathbf{C}_p,\ \kappa,\ \boldsymbol{\beta},\ \theta\},\ \mathbf{G} = \partial\theta/\partial\mathbf{X}, \tag{6a,b}$$

and the constitutive assumptions

$$\psi = \psi(U_1),\ \mathbf{S} = \mathbf{S}(U_1), \tag{7a,b}$$

$$\mathbf{P} = \mathbf{P}(U_1,\ \mathbf{G}),\ \xi = \xi(U_1,\ \mathbf{G}), \tag{7c,d}$$

where $\mathbf{C}_p$ is a symmetric positive definite tensor which denotes plastic deformation, κ denotes isotropic hardening, the tensor $\boldsymbol{\beta}$ denotes directional hardening, θ denotes absolute temperature, $\mathbf{G}$ denotes the temperature gradient, ψ denotes the specific (per unit mass) Helmholtz free energy, $\mathbf{P}$ denotes the entropy flux per unit reference area, and ξ denotes the specific rate of internal entropy supply. The quantities $\mathbf{C}_p$, κ, $\boldsymbol{\beta}$ are defined by evolution equations of the form

$$\mathbf{C}_p = \dot{\mathbf{C}}_p(U_1,\ \mathbf{G}), \tag{8a}$$

$$\dot{\kappa} = \dot{\kappa}(U_1,\ \mathbf{G})\ \dot{\boldsymbol{\beta}} = \dot{\boldsymbol{\beta}}(U_1,\ \mathbf{G}). \tag{8c,d}$$

Furthermore, we note that the directional hardening $\boldsymbol{\beta}$ was

[1]Faculty of Mechanical Engineering, Technion-Israel Institute of Technology, Haifa, Israel 32000.

[2]Mem. ASME.

Manuscript received and accepted by the ASME Applied Mechanics Division, July 20, 1988.

Copyright © 1991 by ASME

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

introduced by Bodner (1985) to model the Bauschinger effect and it is an alternative to kinematic hardening. We also observe that since $\mathbf{C}$, θ, $\mathbf{G}$ are trivially invariant under superposed rigid body motions and the rates (8) are functions of only U_1, $\mathbf{G}$ it follows that $\mathbf{C}_p$, κ, β are trivially invariant under superposed rigid body motions.

Rate-insensitive elastic-plastic materials of the type considered by Green and Naghdi (1965, 1966), but generalized to include directional hardening, may be formulated in terms of a yield function. For this formulation, the constitutive equation (7*d*) and the evolution equations (8) include a linear dependence on the rates

$$\{\dot{\mathbf{C}},\ \dot{\theta}\}. \tag{9}$$

For either the rate-sensitive elastic-viscoplastic material or the rate-insensitive elastic-plastic material it may be shown that the stress $\mathbf{S}$ is determined by the hyperelastic equation

$$\mathbf{S}=2\rho_0(\partial\psi/\partial\mathbf{C}). \tag{10}$$

Following the work of Rubin (1988b) we define the tensors $\mathbf{\Phi}_e$ and $\mathbf{M}_p$ by

$$\mathbf{\Phi}_e=\mathbf{F}\mathbf{M}_p^{-1},\ \mathbf{M}_p=\mathbf{C}_p^{1/2}=\mathbf{M}_p^{T}, \tag{11a, b}$$

where $\mathbf{M}_p$ is uniquely defined in terms of $\mathbf{C}_p$, which itself is defined by the constitutive equation (8*a*). Since the term $\mathbf{M}_p^{-1}$ in equation (11*a*) essentially removes plastic deformation from the deformation gradient $\mathbf{F}$, we may interpret the tensor $\mathbf{\Phi}_e$ as the elastic part of the deformation gradient. Furthermore, let us define

$$\mathbf{C}_e=\mathbf{\Phi}_e^{T}\mathbf{\Phi}_e=\mathbf{M}_p^{-1}\mathbf{C}\mathbf{M}_p^{-1},\quad \mathbf{B}_e=\mathbf{\Phi}_e\mathbf{\Phi}_e^{T}=\mathbf{F}\mathbf{C}_p^{-1}\mathbf{F}^{T}, \tag{12a,b}$$

$$I_{3e}=\det\mathbf{C}_e=I_3/I_{3p},\quad I_{3p}=\det\mathbf{C}_p, \tag{12c,d}$$

where $\mathbf{C}_e$, $\mathbf{B}_e$ are interpreted to be the elastic parts of the tensors $\mathbf{C}$, $\mathbf{B}$, respectively, and I_{3e}, I_{3p} are pure measures of elastic dilatation and plastic dilatation, respectively.

We emphasize that since the quantities in equation (6) are trivially invariant under superposed rigid body motions, it follows directly from equations (11) and (12) that under superposed rigid body motions

$$\mathbf{M}_p^{+}=\mathbf{M}_p,\ \mathbf{\Phi}_e^{+}=\mathbf{Q}\mathbf{\Phi}_e, \tag{13a,b}$$

$$\mathbf{C}_e^{+}=\mathbf{C}_e,\ \mathbf{B}_e^{+}=\mathbf{Q}\mathbf{B}_e\mathbf{Q}^{T}, \tag{13c,d}$$

where a superposed plus denotes the value of the quantity in the superposed configuration and where $\mathbf{Q}(t)$ is an arbitrary proper orthogonal tensor function of time only.

The definitions of equations (11*a*, *b*) should not be confused with definitions related to the separation

$$\mathbf{F}=\mathbf{F}_e\mathbf{F}_p, \tag{14}$$

introduced by Lee (1969) and Lubarda and Lee (1981). Most importantly, we emphasize that $\mathbf{\Phi}_e$ and $\mathbf{M}_p$ are defined uniquely by equations (11) and are not defined by introducing an intermediate stress-free configuration as was done by Lee (1969). It therefore follows that these quantities are not subject to the full invariance restrictions developed by Green and Naghdi (1971) in their analysis of an arbitrariness of the stress-free configuration used by Lee (1969). Furthermore, we note that the tensor $\mathbf{C}_e$ defined in equation (12*a*) is identical to a tensor ($\mathbf{R}_p^{T}\mathbf{C}_e\mathbf{R}_p$ in their notation) introduced by Green and Naghdi (1971).

Using the elastic dilatation I_{3e} in (12*c*) we can define a pure measure of elastic distortion $\mathbf{C}_e'$ by

$$\mathbf{C}_e'=I_{3e}^{-1/3}\mathbf{C}_e,\ \det\mathbf{C}_e'=1. \tag{15a,b}$$

An exact separation of dilatation and distortion of this type was first considerd by Flory (1961) for elasticity and by Simo et al. (1985) for plasticity using different kinematic quantities from those used here. Independently, Rubin (1987a) discussed the significance of a pure measure of elastic distortion in the analysis of metals subjected to high compression.

Since $\mathbf{C}_e'$ and I_{3e} are functions of $\mathbf{C}$ and $\mathbf{C}_p$ it follows that without loss in generality we can formulate constitutive equations in terms of the independent quantities

$$U_2=\{I_3,\ \mathbf{C}_e',\ \mathbf{C}_p,\ \kappa,\ \beta,\ \theta\}, \tag{16}$$

instead of the quantities U_1 in equation (6). Now, for an elastically isotropic material, the Helmholtz free energy ψ can depend on the elastic deformation tensor $\mathbf{C}_e'$ only through its two independent scalar invariants. Alternatively, we could follow the work of Rubin (1987b) and use

$$\beta_1=\mathbf{C}_e'^{-1}\cdot\mathbf{I}=I_{3e}^{1/3}(\mathbf{C}_p\cdot\mathbf{C}^{-1}), \tag{17a}$$

$$\beta_2=\mathbf{C}_e'^{-1}\cdot\mathbf{C}_e'^{-1}=I_{3e}^{2/3}\{\mathbf{C}_p\mathbf{C}^{-1}\cdot\mathbf{C}^{-1}\mathbf{C}_p), \tag{17b}$$

which are two independent scalar invariants of $\mathbf{C}_e'^{-1}$. Then, the constitutive equations for elastically isotropic materials could be formulated in terms of the independent quantities

$$U=\{I_3,\ \beta_1,\ \beta_2,\ \mathbf{C}_p,\ \kappa,\ \beta,\ \theta\}, \tag{18}$$

instead of U_1 or U_2. Using equation (7) with U_1 replaced by U and equation (10) it can be shown that

$$p=-2\rho_0(\partial\psi/\partial I_3)I_3^{1/2} \tag{19a}$$

$$\begin{aligned}\mathbf{S}'=&-2\rho_0(\partial\psi/\partial\beta_1)I_{3e}^{1/3}[\mathbf{C}^{-1}\mathbf{C}_p\mathbf{C}^{-1}-(1/3)(\mathbf{C}_p\cdot\mathbf{C}^{-1})\mathbf{C}^{-1}]\\&-4\rho_0(\partial\psi/\partial\beta_2)I_{3e}^{2/3}[\mathbf{C}^{-1}\mathbf{C}_p\mathbf{C}^{-1}\mathbf{C}_p\mathbf{C}^{-1}\\&-(1/3)(\mathbf{C}_p\mathbf{C}^{-1}\cdot\mathbf{C}^{-1}\mathbf{C}_p)\mathbf{C}^{-1}].\end{aligned} \tag{19b}$$

3 Universal Relations

Using equation (2*c*) and the constitutive equations (19) it may easily be seen that the tensor equation

$$\{\mathbf{C}_p^{-1}\mathbf{C})\mathbf{S}=\mathbf{S}(\mathbf{C}\mathbf{C}_p^{-1}), \tag{20}$$

is satisfied for any functional dependence of ψ on U. In this sense, the equation (20) defines a class of universal relations for elastically isotropic elastic elastic-plastic or elastic-viscoplastic materials. It is important to emphasize that this result holds even when the Helmholtz free energy ψ is an anisotropic function of the tensors $\mathbf{C}_p$ and β and that no restriction need be placed on the functional forms of the evolution equations (8). In particular, the result holds even if plastic deformation is not assumed to be incompressible.

The equation (20) may be rewritten in terms of the Cauchy stress $\mathbf{T}$ and the elastic deformation $\mathbf{B}_e$ by using equation (2*f*) and the definitions (12) to obtain

$$\mathbf{B}_e\mathbf{T}=\mathbf{T}\mathbf{B}_e. \tag{21}$$

It is obvious from the form of equation (21) that for elasticity ($\mathbf{C}_p=\mathbf{I}$, $\mathbf{B}_e=\mathbf{B}$) we obtain Beatty's (1987) equation (3) as a special case. Referring all tensor quantities to a fixed rectangular Cartesian coordinate system with base vectors $\mathbf{e}_i$ ($i=1$, 2, 3), the three nontrivial equations associated with equation (21) may be recorded by replacing the components of $\mathbf{B}$ in equation (7) of Beatty (1987) with those of $\mathbf{B}_e$ to obtain

$$\begin{aligned}&B_{e12}(T_{22}-T_{11})+(B_{e11}-B_{e22})T_{12}\\&\qquad+B_{e13}T_{32}-B_{e32}T_{13}=0,\end{aligned} \tag{22a}$$

$$B_{e13}(T_{33}-T_{11})+(B_{e11}-B_{e33})T_{13}+B_{e12}T_{23}-B_{e23}T_{12}=0, \tag{22b}$$

$$B_{e23}(T_{33}-T_{22})+(B_{e22}-B_{e33})T_{23}+B_{e21}T_{13}-B_{e13}T_{21}=0. \tag{22c}$$

Before closing this section we should mention that in a different context and using different kinematic variables, Lubarda and Lee (1981, p. 36) mentioned the commutativity of certain tensors for elastically isotropic elastic-plastic materials. However, they did not record an equation similar to equation (21) nor did they mention the general nature of the implications of this equation.

4 Pure Shear

As an example, let us derive the universal relations associated with pure shear in the $\mathbf{e}_1-\mathbf{e}_2$ plane for which

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$T_{21} = T_{22} \neq 0, \text{ all other } T_{ij} = 0. \qquad (23a,b)$$

It follows from equation (22) that the universal relations for pure shear yield equations relating the components of $\mathbf{F}$ and $\mathbf{C}_p$ of the forms

$$B_{e11} = B_{e22},\ B_{e13} = 0, \quad B_{e23} = 0. \qquad (24a,b,c)$$

It is important to emphasize that the relations (24) place no restriction on the form of the flow rule (8*a*) for plastic deformation. Therefore, without restricting the form of the flow rule, (8*a*), we cannot obtain more detailed information from the universal relations.

To be more specific let us consider a class of flow rules characterized by the equation

$$\dot{\mathbf{C}}_p = \Gamma \mathbf{A},\ \mathbf{A} = [3/(\mathbf{C}_p^{-1} \cdot \mathbf{C})]\mathbf{C} - \mathbf{C}_p, \qquad (25a,b)$$

where Γ is a function of equation (16) for elastic-viscoplastic response (Rubin, 1987a) and Γ includes a dependence on the quantities (9) for rate-independent elastic-plastic response (Rubin, 1988*b*). For this class of materials, application of shear stress in the $\mathbf{e}_1 - \mathbf{e}_2$ plane will not cause shear deformation in the $\mathbf{e}_1 - \mathbf{e}_3$ or $\mathbf{e}_2 - \mathbf{e}_3$ planes so the deformation may be represented by

$$x_1 = a(X_1 + \gamma X_2),\ x_2 = bX_2,\ x_3 = cX_3, \qquad (26a,b,c)$$

where γ is a measure of shear and a, b, c are scalars to be determined. Taking $\mathbf{C}_p = \mathbf{I}$ initially, it follows from equations (2*g*), (19*b*), (25), and (26) that

$$C_{p13} = 0, \quad C_{p23} = 0, \qquad (27a,b)$$

$$T_{13} = 0 \quad T_{23} = 0, \qquad (27c,d)$$

for any values of γ, a, b, c.

Since the deformation is homogeneous the state of pure shear can be obtained by solving the three conditions

$$T_{11} = 0 \quad T_{22} = 0, \quad T_{33} = 0 \qquad (28a,b,c)$$

for the scalars a, b, c as functions of the shear γ. After some algebraic manipulation it may be shown, using equations (12*b*) and (26), that the universal relation (24*a*) requires

$$a^2(C_{p22} + \gamma^2 C_{p11} - 2\gamma C_{p12}) = b^2 C_{p11}, \qquad (29)$$

and the universal relations (24*b*, *c*) are automatically satisfied by the conditions (27*a*, *b*). Furthermore, we note that the simple expression (29) is not independent of the conditions (28) so that it can be used to replace one of them. Since the expressions for the stresses T_{ij} can be quite complicated this could result in considerable analytical simplification.

Finally, we mention that for elastic response ($\mathbf{C}_p = \mathbf{I}$) equation (29) yields the universal relation

$$a^2(1 + \gamma^2) = b^2, \qquad (30)$$

which was deduced by Rubin (1988a) in the elastic analysis of pure shear.

References

Beatty, M. F., 1987, "A Class of Universal Relations In Isotropic Elasticity Theory," *J. Elasticity*, Vol. 17, pp. 113–121.

Bodner, S. R., 1985, "Evolution Equations for Anisotropic Hardening and Damage of Elastic-Viscoplastic Materials," *Plasticity Today—Modeling, Methods and Applications*, A. Sawczuk, and G. Bianchi, eds., Elsevier, London, pp. 471–482.

Flory, P. J., 1961, "Thermodynamic Relations For High Elastic Materials," *Trans. Faraday Soc.*, Vol. 57, pp. 829–838.

Green, A. E., and Naghdi, P. M., 1965, "A General Theory of An Elastic-Plastic Continuum," *Arch. Rat. Mech. Anal.*, Vol. 18, pp. 251–281.

Green, A. E., and Naghdi, P. M., 1966, "A Thermodynamic Development of Elastic-Plastic Continua," *Proceedings of the IUTAM Symposia*, Vienna, June 22–28, 1966, H. Parkus, and L. I. Sedov. eds., Springer-Verlag, New York, pp. 117–131.

Green, A. E., and Naghdi, P. M., 1971, "Some Remarks On Elastic-Plastic Deformation At Finite Strain," *Int. J. Engng. Sci.*, Vol. 9, pp. 1219–1229.

Green, A. E., and Naghdi, P. M., 1977, "On Thermodynamics And The Nature Of the Second Law," *Proc. Royal Soc. Lond.*, Vol. A357, pp. 253–270.

Green, A. E., and Naghdi, P. M., 1978, "The Second Law Of Thermodynamics And Cyclic Processes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 487–492.

Lee, E. H., 1969, "Elastic-Plastic Deformation at Finite Strains," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 1–6.

Lubarda, V. A., and Lee, E. H., 1981, "A Correct Definition Of Elastic And Plastic Deformation And Its Computational Significance," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 35–40.

Rubin, M. B., 1986, "An Elastic-Viscoplastic Model for Large Deformation," *Int. J. Engng. Sci.*, Vol. 24, pp. 1083–1095.

Rubin, M. B., 1987a, "An Elastic-Viscoplastic Model for Metals Subjected to High Compression," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 532–538.

Rubin, M. B., 1987b, "An Elastic-Viscoplastic Model Exhibiting Continuity of Solid And Fluid States," *Int. J. Engng. Sci.*, Vol. 25, pp. 1175–1191.

Rubin, M. B., 1988a, "The Significance of Pure Measures of Distortion in Nonlinear Elasticity With Reference To the Pointing Problem," *J. Elasticity*, Vol. 20, pp. 53–64.

Rubin, M. B., 1988b, "The Significance of Nonhydrostatic Distortional Strain in Anisotropic Elastic-Plastic Metals," submitted for publication.

Simo, J. C., Taylor, R. I., and Pister, K. S., 1985, "Variational And Projection Methods for The Volume Constraint in Finite Deformation Elasto-Plasticity," *Computer Meth. Appl. Mech. Engng.*, Vol. 51, pp. 177–208.

Truedell, C., and Noll, W., 1965, *The Non-Linear Field Theories of Mechanics*, Encyclopedia of Physics, S. Flugge, ed., Vol. III/3.

Response of a Crack to Transient Concentrated Line Forces

Chen Baoxing[3] and Zhang Xiangzhou[3]

1 Introduction

Of importance in structural analysis is the response of a crack to a time-dependent stress field. In two-dimensional cases the response has been considered by many investigators. However, from the published references it is seen that the study appears to be confined within the scope of plane incident waves or uniform impact tractions. When an elastic medium is stressed by some transient concentrated line forces, curved wave fronts will be generated to impinge cracks. Or, in other words, the crack surfaces will be in motion by the application of surface tractions—nonuniform with respect both to space and time. In spite of the fact that the corresponding study in a static case has already been well established (Rooke and Cartwright, 1976), until now it seems no investigation has treated the subject of response of cracks to transient concentrated line forces.

In this Note, the response of a crack contained in an infinite medium to transient concentrated line forces is studied. The problem is first reduced to that of a crack under the action of certain nonuniform tractions applied on its surfaces. Laplace and Fourier transforms are then used to formulate the mixed boundary value problem finally to a Fredholm integral equation of the second kind in Laplace transform domain. When applying numerical inversion with some specially designed procedures, the values of dynamic stress intensity factors are obtained and presented in diagrams. Some conclusions regarding the variation of the dynamic stress intensity factors, with respect to time, crack geometries, and material constants are drawn in the concluding portion of this Note. It is found that the points of application of the concentrated line forces, which induce the highest value of the dynamic stress intensity factors, are not on the crack surfaces but at some distance away from the crack.

[3]Department of Aircraft Engineering, Northwestern Polytechnical University, Xi'an, 710072, China.

Manuscript received by the ASME Applied Mechanics Division June 6, 1989; final revision, Jan. 10, 1990.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$T_{21} = T_{22} \neq 0, \text{ all other } T_{ij} = 0. \qquad (23a,b)$$

It follows from equation (22) that the universal relations for pure shear yield equations relating the components of $\mathbf{F}$ and $\mathbf{C}_p$ of the forms

$$B_{e11} = B_{e22},\ B_{e13} = 0, \quad B_{e23} = 0. \qquad (24a,b,c)$$

It is important to emphasize that the relations (24) place no restriction on the form of the flow rule (8*a*) for plastic deformation. Therefore, without restricting the form of the flow rule, (8*a*), we cannot obtain more detailed information from the universal relations.

To be more specific let us consider a class of flow rules characterized by the equation

$$\dot{\mathbf{C}}_p = \Gamma \mathbf{A},\ \mathbf{A} = [3/(\mathbf{C}_p^{-1} \cdot \mathbf{C})]\mathbf{C} - \mathbf{C}_p, \qquad (25a,b)$$

where Γ is a function of equation (16) for elastic-viscoplastic response (Rubin, 1987a) and Γ includes a dependence on the quantities (9) for rate-independent elastic-plastic response (Rubin, 1988*b*). For this class of materials, application of shear stress in the $\mathbf{e}_1 - \mathbf{e}_2$ plane will not cause shear deformation in the $\mathbf{e}_1 - \mathbf{e}_3$ or $\mathbf{e}_2 - \mathbf{e}_3$ planes so the deformation may be represented by

$$x_1 = a(X_1 + \gamma X_2),\ x_2 = bX_2,\ x_3 = cX_3, \qquad (26a,b,c)$$

where γ is a measure of shear and a, b, c are scalars to be determined. Taking $\mathbf{C}_p = \mathbf{I}$ initially, it follows from equations (2*g*), (19*b*), (25), and (26) that

$$C_{p13} = 0, \quad C_{p23} = 0, \qquad (27a,b)$$

$$T_{13} = 0 \quad T_{23} = 0, \qquad (27c,d)$$

for any values of γ, a, b, c.

Since the deformation is homogeneous the state of pure shear can be obtained by solving the three conditions

$$T_{11} = 0 \quad T_{22} = 0, \quad T_{33} = 0 \qquad (28a,b,c)$$

for the scalars a, b, c as functions of the shear γ. After some algebraic manipulation it may be shown, using equations (12*b*) and (26), that the universal relation (24*a*) requires

$$a^2(C_{p22} + \gamma^2 C_{p11} - 2\gamma C_{p12}) = b^2 C_{p11}, \qquad (29)$$

and the universal relations (24*b*, *c*) are automatically satisfied by the conditions (27*a*, *b*). Furthermore, we note that the simple expression (29) is not independent of the conditions (28) so that it can be used to replace one of them. Since the expressions for the stresses T_{ij} can be quite complicated this could result in considerable analytical simplification.

Finally, we mention that for elastic response ($\mathbf{C}_p = \mathbf{I}$) equation (29) yields the universal relation

$$a^2(1 + \gamma^2) = b^2, \qquad (30)$$

which was deduced by Rubin (1988a) in the elastic analysis of pure shear.

References

Beatty, M. F., 1987, "A Class of Universal Relations In Isotropic Elasticity Theory," *J. Elasticity*, Vol. 17, pp. 113–121.

Bodner, S. R., 1985, "Evolution Equations for Anisotropic Hardening and Damage of Elastic-Viscoplastic Materials," *Plasticity Today—Modeling, Methods and Applications*, A. Sawczuk, and G. Bianchi, eds., Elsevier, London, pp. 471–482.

Flory, P. J., 1961, "Thermodynamic Relations For High Elastic Materials," *Trans. Faraday Soc.*, Vol. 57, pp. 829–838.

Green, A. E., and Naghdi, P. M., 1965, "A General Theory of An Elastic-Plastic Continuum," *Arch. Rat. Mech. Anal.*, Vol. 18, pp. 251–281.

Green, A. E., and Naghdi, P. M., 1966, "A Thermodynamic Development of Elastic-Plastic Continua," *Proceedings of the IUTAM Symposia*, Vienna, June 22–28, 1966, H. Parkus, and L. I. Sedov. eds., Springer-Verlag, New York, pp. 117–131.

Green, A. E., and Naghdi, P. M., 1971, "Some Remarks On Elastic-Plastic Deformation At Finite Strain," *Int. J. Engng. Sci.*, Vol. 9, pp. 1219–1229.

Green, A. E., and Naghdi, P. M., 1977, "On Thermodynamics And The Nature Of the Second Law," *Proc. Royal Soc. Lond.*, Vol. A357, pp. 253–270.

Green, A. E., and Naghdi, P. M., 1978, "The Second Law Of Thermodynamics And Cyclic Processes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 487–492.

Lee, E. H., 1969, "Elastic-Plastic Deformation at Finite Strains," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 1–6.

Lubarda, V. A., and Lee, E. H., 1981, "A Correct Definition Of Elastic And Plastic Deformation And Its Computational Significance," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 35–40.

Rubin, M. B., 1986, "An Elastic-Viscoplastic Model for Large Deformation," *Int. J. Engng. Sci.*, Vol. 24, pp. 1083–1095.

Rubin, M. B., 1987a, "An Elastic-Viscoplastic Model for Metals Subjected to High Compression," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 532–538.

Rubin, M. B., 1987b, "An Elastic-Viscoplastic Model Exhibiting Continuity of Solid And Fluid States," *Int. J. Engng. Sci.*, Vol. 25, pp. 1175–1191.

Rubin, M. B., 1988a, "The Significance of Pure Measures of Distortion in Nonlinear Elasticity With Reference To the Pointing Problem," *J. Elasticity*, Vol. 20, pp. 53–64.

Rubin, M. B., 1988b, "The Significance of Nonhydrostatic Distortional Strain in Anisotropic Elastic-Plastic Metals," submitted for publication.

Simo, J. C., Taylor, R. I., and Pister, K. S., 1985, "Variational And Projection Methods for The Volume Constraint in Finite Deformation Elasto-Plasticity," *Computer Meth. Appl. Mech. Engng.*, Vol. 51, pp. 177–208.

Truedell, C., and Noll, W., 1965, *The Non-Linear Field Theories of Mechanics*, Encyclopedia of Physics, S. Flugge, ed., Vol. III/3.

Response of a Crack to Transient Concentrated Line Forces

Chen Baoxing[3] and Zhang Xiangzhou[3]

1 Introduction

Of importance in structural analysis is the response of a crack to a time-dependent stress field. In two-dimensional cases the response has been considered by many investigators. However, from the published references it is seen that the study appears to be confined within the scope of plane incident waves or uniform impact tractions. When an elastic medium is stressed by some transient concentrated line forces, curved wave fronts will be generated to impinge cracks. Or, in other words, the crack surfaces will be in motion by the application of surface tractions—nonuniform with respect both to space and time. In spite of the fact that the corresponding study in a static case has already been well established (Rooke and Cartwright, 1976), until now it seems no investigation has treated the subject of response of cracks to transient concentrated line forces.

In this Note, the response of a crack contained in an infinite medium to transient concentrated line forces is studied. The problem is first reduced to that of a crack under the action of certain nonuniform tractions applied on its surfaces. Laplace and Fourier transforms are then used to formulate the mixed boundary value problem finally to a Fredholm integral equation of the second kind in Laplace transform domain. When applying numerical inversion with some specially designed procedures, the values of dynamic stress intensity factors are obtained and presented in diagrams. Some conclusions regarding the variation of the dynamic stress intensity factors, with respect to time, crack geometries, and material constants are drawn in the concluding portion of this Note. It is found that the points of application of the concentrated line forces, which induce the highest value of the dynamic stress intensity factors, are not on the crack surfaces but at some distance away from the crack.

[3]Department of Aircraft Engineering, Northwestern Polytechnical University, Xi'an, 710072, China.

Manuscript received by the ASME Applied Mechanics Division June 6, 1989; final revision, Jan. 10, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

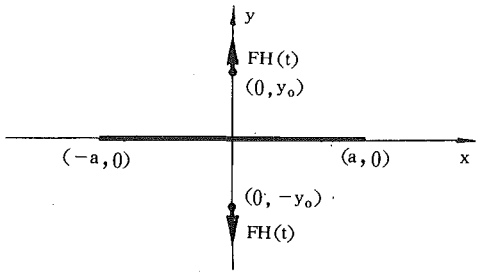


Fig. 1 The crack and the concentrated forces

2 Problem and Solution

An infinite elastic medium containing a crack defined by $|x| \le a$, $y=0$ is depicted in Fig. 1. The medium is under the state of plane strain caused by the action of a pair of transient concentrated line forces $FH(t)$ as shown in the figure, with $H(t)$ being the Heaviside unit function. It is required to find the dynamic stress intensity factors at the crack tips. By the principle of superposition we see that the solution to this problem can be obtained by summing the solutions to the following two problems: (a) the infinite medium under the action of the concentrated line forces $FH(t)$ (see Fig. 1, but the crack should be removed) and (*b*) the cracked infinite medium under the action of normal tractions on the crack surfaces, equal and opposite to those found from solution (a) (see Fig. 1, but the concentrated forces should be removed).

For plane elastodynamic problems, the displacement and stress components can be expressed in terms of two scalar potentials, $\varphi(x,y,t)$ and $\psi(x,y,t)$, which satisfying the well-known wave equations (Achenbach, 1973). The solution of problem (a) can be obtained by the use of available singular solutions in elastodynamics (Achenbach, 1973). The solution of the problem (b) should be determined from the wave equations (Achenbach, 1973) and the following conditions:

$$u(x,y,0)=v(x,y,0)=\dot{u}(x,y,0)=\dot{v}(x,y,0)=0, \tag{1}$$

$$\sigma_{xy}(x,0,t)=0, \quad -\infty<x<\infty, \quad 0<t<\infty, \tag{2}$$

$$v(x,0,t)=0, \quad a\le|x|<\infty, \quad 0<t<\infty, \tag{3a}$$

$$\sigma_{yy}(x,0,t)=-\sigma^0_{yy}(x,0,t), |x|<a, 0<t<\infty, \tag{3b}$$

$$\lim_{x^2+y^2\to\infty}[u(x,y,t),\ v(x,y,t),\ \ldots]=0, \tag{4}$$

where $\sigma^0_{yy}(x,0,t)$ stands for the normal stress distribution on the line $y=0$ in the problem (a). Using a procedure similar to that in Chen and Sih (1977), the solutions in Laplace transform domain are obtained:

$$\varphi^*(x,y,p)=\frac{1}{2\pi}\int_{-\infty}^{\infty} -(s^2+p^2/2C_2^2)G(s,p)\exp(-\eta_1 y-isx)\,ds, \tag{5a}$$

$$\psi^*(x,y,p)=\frac{1}{2\pi}\int_{-\infty}^{\infty} is\,\eta_1 G(s,p)\exp(-\eta_2 y-isx)\,ds, \tag{5b}$$

where $\eta_i=\{(p/C_i)^2+s^2\}^{1/2}$, $(i=1,2)$, and C_1, C_2, being the dilatational and shear wave velocity of the material, respectively, p and s are transform variables of Laplace and Fourier transforms, whereas φ^* and ψ^* denote the Laplace transform of function φ and ψ, respectively. The unknown function $G(s,p)$ can be written in the following form

$$\eta_1 G(s,p)= -\{aF/[\mu p(p^2/C_1^2-p^2/C_2^2)]\}\int_0^1 \sqrt{\tau}\,D(\tau,p)\,J_0(as\tau)\,d\tau, \tag{6}$$

where $J_0(x)$ is the Bessel function of order zero, and $D(\tau,p)$ is prescribed by the following Fredholm integral equation of the second kind

$$D(\tau,p)+\int_0^1 D(\gamma,p)K(\gamma,\tau,p)d\gamma=B(\tau,p), \tag{7}$$

where

$$K(\gamma,\tau,p)=-\sqrt{\gamma\tau}\int_0^\infty sw(s/a,p)J_0(s\gamma)J_0(s\tau)ds,$$

$$B(\tau,p)=2\sqrt{\tau}\int_0^\tau [ap\,\sigma^{0*}_{yy}(a\xi,0,p)/(F\sqrt{\tau^2-\xi^2})]d\xi,$$

$$w(s,p)=1+[(2s^2+p^2/C_2^2)^2-4s^2\eta_1\eta_2]/[s\eta_1 2(p^2/C_1^2-p^2/C_2^2)].$$

When $D(\tau,p)$ is solved from equation (7), then from equation (5) $\varphi^*(x,y,p)$ and $\psi^*(x,y,p)$ in the Laplace transform domain are obtained. Following the Laplace inverse transform, the displacements and stresses can be determined.

The dynamic stress intensity factor is of foremost significance in engineering application. Expressions for the dynamic stress intensity factor can be derived as follows: When substituting the expressions of φ^* and ψ^* (equation (5)) into the transformed equation of the stress components, it is not difficult to obtain

$$\sigma^*_{yy}(x,0,p)=[(F/\pi a)x/\sqrt{x^2-a^2}\,]D(1,p)/p + \text{nonsingular terms}. \tag{8}$$

The dynamic stress intensity factors can therefore be obtained in Laplace transform domain as follows:

$$K_1^*(p)=\lim_{x\to a^+}\sqrt{2(x-a)}\,\sigma^*_{yy}(x,0,p)=(F/\pi\sqrt{a})\,D(1,p)/p. \tag{9}$$

Applying the Laplace inverse transform to equation (9) then yields

$$K_1(t)=\frac{1}{2\pi i}\int_{Br} K_1^*(p)\exp(pt)\,dp, \tag{10}$$

where Br stands for a Bromwich contour.

3 Evaluation of Dynamic Stress Intensity Factors

Numerical evaluation of dynamic stress intensity factors can now be carried out via the following steps:

(1) determination of Laplace transform of $\sigma^0_{yy}(x,0,t)$, and

(2) determination of the kernel function $K(\gamma,\tau,p)$.

(The details of the two steps are conventional and are omitted here.)

(3) evaluation of the nonhomogeneous term of the Fredholm integral equation.

For this purpose, use is made of the substitution $\xi=\tau\sin\theta$. The singular integral of $B(\tau,p)$ in equation (7) is then reduced to a normal definite integral

$$B(\tau,p)=2\sqrt{\tau}\int_0^{\pi/2} ap\,\sigma^{0*}_{yy}(a\tau\sin\theta,0,p)\,d\theta/F, \tag{11a}$$

which can be integrated by some conventional numerical quadrature techniques.

For $y_0/a=0$, that is, when the transient concentrated line forces apply on the crack surfaces, $B(\tau,p)$ takes the following simple form

$$B(\tau,p)=1/\sqrt{\tau}. \tag{11b}$$

(4) numerical solution of the Fredholm integral equation.

To obtain dynamic stress intensity factor $K_1^*(p)$ in the Laplace transform domain it is necessary to evaluate $D(1,p)$ in advance. For this purpose discretize the Fredholm equation in the following system of linear algebraic equations

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

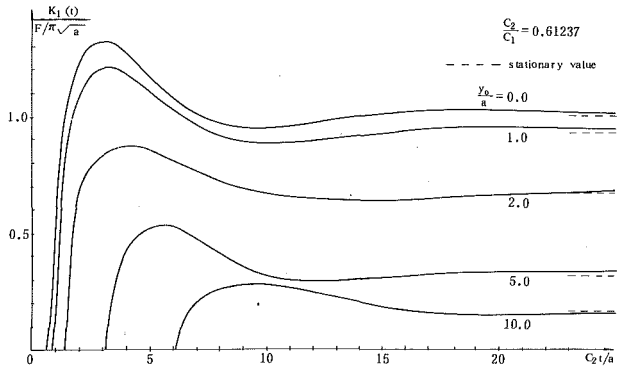


Fig. 2 Values of normalized dynamic stress intensity factors

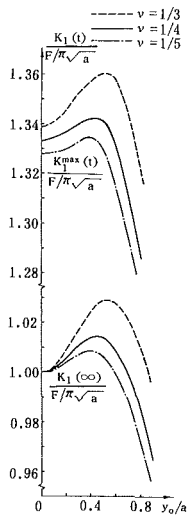


Fig. 3 Values of normalized dynamic and stationary stress intensity factors

$$D(\tau_i,p)+\sum_{j=1}^{n} D(\gamma_j,p)\,K(\gamma_j,\tau_i,p)\Delta\gamma_j=B(\tau_i,p),$$
$$i=1,2,\ldots,n, \tag{12}$$

where $\Delta\gamma_j=\gamma_j-\gamma_{j-1}$.

For a certain value of p, solution of the algebraic equation system gives the value of $D(\tau, p)$ at a discrete set of points $\tau=\tau_i$ $(i=1,2,\ldots,n)$. The value of $D(1, p)$ can be obtained from a quadratic extrapolation of $D(\tau_i, p)$ based on the last three points $(\tau_i=\tau_{n-2}, \tau_{n-1}, \tau_n)$.

(5) determination of dynamic stress intensity factors.

Numerical inversion of Laplace transform has been carried out to obtain the values of dynamic stress intensity factors (Miller and Guy, 1966). To this end, $K_1(t)$ is written in the following form

$$K_1(t)\approx\frac{F}{\pi\sqrt{a}}\sum_{n=0}^{N} C_nP_n^{(0,\beta)}[2\exp(-\delta t)-1], \tag{13}$$

where $P_n^{(0,\beta)}(x)$ is Jacobi polynomial of degree n and C_n can be determined from the following algebraic system

$$\frac{D[1,(\beta+1+k)\delta]}{\beta+1+k}=\sum_{m=0}^{k}\frac{k(k-1)\ldots[k-(m-1)]}{(k+\beta+1)(k+\beta+2)\ldots(k+\beta+1+m)}C_m, \tag{14}$$

with β and δ being parameters used in the numerical computation.

Generally speaking, there is no universally optimum way of selecting these values. In order to get better results the parameters β and δ should be selected such that $K_1(t)$ can be best described within a particular range of time t, including $t\to\infty$, the static case. In this Note, $\beta=0.0$, $C_2/\delta a=4.0\sim13.0$ and $N=4\sim6$ are selected.

4 Numerical Results and Conclusions

Numerical evaluation of the dynamic stress intensity factors has been carried out for a number of geometry parameters and three material parameters $C_2/C_1=0.61237, 0.57735, 0.5$. These values of C_2/C_1 correspond to $\nu=1/5, 1/4, 1/3$, respectively, with ν being the Poisson's ratio of the material.

Because of the limit of length, only one group of the curves of the normalized dynamic stress intensity factors for $C_2/C_1=0.61237$ are shown in Fig. 2. The result shown in Fig. 2 demonstrates a general feature of the dynamic stress intensity factors common to various materials and locations of the concentrated line forces. For small value of t, the dynamic stress intensity factors are zero and then increase rapidly from $t=\sqrt{a^2+y_0^2}/C_1$, corresponding to the instant when the dilatational wave first arrives at the crack tips. After reaching a maximum, the dynamic stress intensity factors begin to decrease and then oscillate around the corresponding stationary values.

The dashed lines in Fig. 2 stand for the corresponding stationary stress intensity factors, that is, the value of $K_1(t)$ when t tends to infinity. The stationary values $K_1(\infty)$ are obtained from the following formula (Rooke and Cartwright, 1976).

$$K_1(\infty)=\{1/\sqrt{1+y_0^2/a^2}+[(y_0^2/a^2)/(1+y_0^2/a^2)^{3/2}]/[2(1-\nu)]\}\cdot F/\pi\sqrt{a}. \tag{15}$$

It can be shown rigorously that the maximum value of $K_1(\infty)$

$$K_1^{\max}(\infty)=[(1-2\nu/3)^{3/2}/(1-\nu)]\cdot F/\pi\sqrt{a} \tag{16}$$

reaches at the point $y_0/a=\sqrt{2\nu/(3-2\nu)}$, which is some distance away from the crack except in the case $\nu=0$ for which $y_0=0$. For $\nu=1/5, 1/4, 1/3$, $y_0/a=0.3922, 0.4472$, and 0.5345, respectively. Therefore, the points of application of the concentrated line forces, which induce the highest value of the stationary stress intensity factors, are at some distance away from the crack rather than on the crack surfaces.

For transient concentrated line forces numerical results show a similar situation. This can be seen from Fig. 3 at a point near the point $(0, \sqrt{2\nu/(3-2\nu)}\,a)$ the dynamic stress intensity factor reaches its highest value. This conclusion is of practical value in engineering. The concentrated line forces should not apply at the "critical point," to avoid excessive large value of stress intensity factors.

It can be seen from Figs. 2–3 that for most cases the ratio $K_1^{\max}(t)/K_1(\infty)=1.30\sim1.34$. For large values of y_0/a the ratio enlarges. However, this would be of little consequence since for large y_0/a both $K_1^{\max}(t)$ and $K_1(\infty)$ are relatively small.

References

Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North Holland Publishing Company, Amsterdam.

Chen, E. P., and Sih, G. C., 1977, "Transient Response of Cracks to Impact Loads," *Elastodynamic Crack Problem*, (Mechanics of Fracture), Vol. 4, G.C. Sih, ed., Noordhoff International Publishing Company, Leyden.

Miller, M. K., and Guy, W. T., 1966, "Numerical Inversion of the Laplace Transform by Use of Jacobi Polynomials," *SIAM J. Numer. Anal*, Vol. 3, pp. 624–635.

Rooke, D. P., and Cartwright, D. J., 1976, *Compendium of Stress Intensity Factors*, Her Majesty's Stationary Office.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The Shear Modulus of Foamlike Structures

D. Stamenović[4]

Introduction

In modeling elastic behavior of liquid foams during small deformation, many authors have employed a microstructural approach (cf., Stamenović and Wilson, 1984; Kraynik and Hansen, 1986; Budiansky and Kimmel, 1990). In those studies, foam microstructure is pictured as a network of interconnected liquid films that provide surface tension (γ) at the air-liquid interface. The authors assume that surface forces equilibrate between adjacent films and hence the value of γ is the same on all faces and a function of the total interfacial surface area (S), $\gamma=\gamma(S)$. However, in each of those studies either an idealized microstructural geometry is developed (Kraynik and Hansen, 1986; Budiansky and Kimmel, 1990) or affine microstructural strain assigned (Stamenović and Wilson, 1984). The shear modulus (μ) predicted from those models turns out to be a constant fraction of the overpressure of the air trapped inside foam (ΔP), independent of the surface tension-surface area properties of the surface active foaming agent. However, there is an appreciable quantitative discrepancy between the values predicted from different models, as well as between the predicted values and the one observed from rheometric tests on foams and emulsions (cf., Stamenović and Wilson, 1984; Princen and Kiss, 1986). These indicate the significance of geometry and kinematics of the microstructure on shear behavior of foam. Our purpose here is to demonstrate that one can obtain μ that is qualitatively consistent with the previous findings, assuming no particular geometry or kinematics of the microstructure.

Derivation

Like in the previous studies, we regard foam as an elastic isotropic material with the reference configuration at the state of uniform prestress ΔP. We also assume that γ is equal on the adjacent surfaces. To calculate μ, we consider foam during a small simple shear, given by shear strain θ. Let τ denote the corresponding shear stress. Then, from the principle of virtual work, we have that

$$\tau\delta\theta=(1/V_0)\gamma(S)\delta S=(1/V_0)\gamma(S)(\partial S/\partial\theta)\delta\theta, \tag{1}$$

where V_0 is the volume occupied by foam at the reference configuration. From the power series expansion of $\gamma(S)$ and S around the reference configuration, we obtained the following:

$$\gamma(S)=\gamma_0+(d\gamma/dS)_0(S-S_0)+(1/2)(d^2\gamma/dS^2)_0(S-S_0)^2+O(|S-S_0|^3), \tag{2}$$

$$S=S_0+(\partial S/\partial\theta)_0\theta+(1/2)(\partial^2 S/\partial\theta^2)_0\theta^2+O(|\theta|^3), \tag{3}$$

where subscript 0 denotes the function evaluated at the reference configuration. Note that the expansion given by equation (3) assumes, in effect, that there exists a relationship between microscopic strain and macroscopic strain without explicitly specifying what that relationship is. In this respect, our approach is more general than the previous ones in which this relationship was determined from the specific geometry or kinematics of the microstructure. Combining equations (1)–(3), we obtained the following:

$$\tau=(\gamma_0/V_0)(\partial S/\partial\theta)_0+(\gamma_0/V_0)(\partial^2 S/\partial\theta^2)_0\theta+(1/V_0)(d\gamma/dS)_0(\partial S/\partial\theta)_0^2\theta+O(|\theta|^2). \tag{4}$$

Since $\theta=0$ implies $\tau=0$, it follows from equation (4) that $(\partial S/\partial\theta)_0=0$. Thus,

$$\tau=(\gamma_0/V_0)(\partial^2 S/\partial\theta^2)_0\theta+O(|\theta|^2). \tag{5}$$

We identified μ with the coefficient of the linear term in θ in equation (5), i.e.,

$$\mu=(\gamma_0/V_0)(\partial^2 S/\partial\theta^2)_0. \tag{6}$$

It is shown that for a three-dimensional foamlike structure, $\Delta P=2\gamma_0 S_0/3V_0$, independent of a particular microstructural geometry (cf., Stamenović and Wilson, 1984; Princen and Kiss, 1986; Budiansky and Kimmel, 1990). Taking this into account, equation (6) can be rewritten as follows:

$$\mu=\frac{3}{2}\frac{\partial^2(S/S_0)}{\partial\theta^2}\Big|_0\Delta P. \tag{7}$$

For a two-dimensional foamlike structure, $\Delta P=\gamma_0 S_0/2A_0$, where A_0 is the reference surface area occupied by foam and S_0 and γ_0 have notions of the total length of the microstructural line elements and force in the line elements, respectively, at the reference configuration (cf., Kraynik and Hansen, 1986). In this case, equation (6) can be rewritten as follows.

$$\mu=2\frac{\partial^2(S/S_0)}{\partial\theta^2}\Big|_0\Delta P. \tag{8}$$

Concluding Remarks

The most attractive feature of the above derivation is that it yielded a simple, physically plausible, and mathematically transparent result, starting from the principle of virtual work. It required only one assumption, i.e., that γ on the adjacent surfaces is equal, whereas in the previous models, geometry or kinematics of the microstructure are also specified. This result is instructive from the several aspects given next.

First, equations (7) and (8) indicate that mechanical behavior of foamlike structures during small shear deformation is entirely determined by the pressure in the gas in the cells and by the geometrical features of the microstructure. It does not depend on the surface tension-surface area properties of the surface active foaming agent. This is consistent with the behavior predicted from earlier models and the observed behavior of soap foam.

Second, by calculating $\partial^2(S/S_0)/\partial\theta^2|_0$ from the models of Stamenović and Wilson (1984) and Kraynik and Hansen (1986), and substituting those values into equations (7) and (8), respectively, we obtained values of μ that were consistent with the values obtained from those models.

Third, elastic stability requirements dictate $\mu>0$ which in turn, according to equation (6), implies $(\partial^2 S/\partial\theta^2)_0>0$. This, together with the fact that $(\partial S/\partial\theta)_0=0$, imply that S attains its minimum at the reference configuration. This is consistent with the isoperimetric character of the liquid film, i.e., it assumes the minimal surface area at equilibrium.

Fourth, the foregoing result could be useful in studies of lung elasticity. Lung parenchyma is a foamlike structure with a considerable surface energy. The parenchymal shear modulus appears to be a constant fraction of transpulmonary gas pressure for a wide range of pressures (Lai-Fook, 1979). However, besides surface forces, the parenchymal tissue system contributes to parenchymal mechanical behavior. It is not yet fully understood which of those two systems has a predominant role in determining the parenchymal shear modulus. Our result suggests that more information regarding this problem can be obtained by observing the change of the parenchymal interfacial surface area during small shear deformation. If surface network of lung parenchyma were indeed

[4]Department of Biomedical Engineering, Boston University, Boston, MA 02215.

Manuscript received and accepted by the ASME Applied Mechanics Division, Oct. 4, 1989.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

mechanically analogous to liquid foams, the change of parenchymal interfacial surface area with macroscopic shear strain should have a similar pattern as the change of foam surface area with the macroscopic shear strain.

Acknowledgment

The work was supported by National Heart, Lung, and Blood Institute, Grant HL 36958, and the Whitaker Foundation Grant.

References

Budiansky, B., and Kimmel, E., 1990, "On the Shear Modulus of Polyhedron-Cell Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp.

Kraynik, A. M., and Hansen, M. G., 1986, "Foam Emulsion Rheology: A Quasistatic Model for Large Deformations of Spatially-Periodic Cells," *Journal of Rheology*, Vol. 30, pp. 409–439.

Lai-Fook, S. J., 1979, "The Elastic Constants of Lung Parenchyma: The Effect of Pressure-Volume Hysteresis on the Behavior of Blood Vessels," *Journal of Biomechanics*, Vol. 12, pp. 757–764.

Princen, H. M., and Kiss, A. D., 1986, "Rheology of Foams and Highly Concentrated Emulsions. III. Static Shear Modulus," *Journal of Colloid and Interface Science*, Vol. 112, pp. 427–437.

Stamenović, D., and Wilson, T. A., 1984, "The Shear Modulus of Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 229–231.

On the Shear Modulus of Polyhedron-Cell Liquid Foam

B. Budiansky[5] and E. Kimmel[6]

Introduction

We consider liquid foam to be a cluster of closed cells bounded by thin, liquid walls, containing trapped gas. At a sufficiently high gas fraction, the cells are contiguous polyhedra of various sizes, shapes, and orientations (Aubert, 1986).

The pressure P_i of the trapped gas in each cell is slightly higher than the pressure P_e external to the foam. The overpressure, $P_c = P_i - P_e$, the so-called capillary pressure (Bikerman, 1973) is presumed to be balanced by surface tension forces in the cell walls.

For small external loading, in the absence of dissipation, the liquid foam may be idealized as an elastic isotropic material. Since the surface tension and the capillary pressure are constrained by equilibrium, the mechanical properties of the foam should depend only on the capillary pressure and the compressibility of the trapped gas.

By means of a Deer rheometer, Stamenovic and Wilson (1984) measured the elastic response of foam to an applied torque. Their measurements gave values of the ratio of the shear modulus G to capillary pressure P_c that ranged between .24 and .41, with an average value $G/P_c = .33$.

The somewhat larger result $G/P_c = .4$ was calculated theoretically by Stamenovic and Wilson on the basis of the assumption that the foam is a collection of randomly-oriented plates that deform in an affine manner.[7] However, recent studies of lung elasticity (Kimmel, Kamm, and Shapiro, 1987; Budiansky and Kimmel, 1987; Kimmel and Budiansky, 1989) suggest that the affine-deformation assumption is unduly constraining when applied to foam-like polyhedral assemblages. In this note we modify the analysis of Kimmel and Budiansky (1989) to derive the elastic moduli of foam.

Analysis

We first write the equilibrium connection between the capillary pressure P_c and the surface tension Γ per unit length that acts at each of the gas-liquid interfaces on each cell wall. By the principal of virtual work

$$P_c \Sigma \delta V_i = \Gamma \Sigma \delta S_i, \tag{1}$$

where the summations are over all the cells of the foam, V_i is the volume of the ith cell, and S_i is its boundary area. For self-similar (i.e., affine) virtual displacements, we have

$$\frac{\delta V_i}{V_i} = \frac{3}{2}\frac{\delta S_i}{S_i} \tag{2}$$

whence equation (1) implies

$$D \equiv \frac{\Gamma S_o}{P_c V_o} = \frac{3}{2} \tag{3}$$

where V_o and S_o are the average cell volume and boundary surface area. This relation has been derived repeatedly in various ways (Ross, 1969; Hoppin and Hildebrandt, 1979; Stamenovic and Wilson, 1984).

Next, we invoke the dodecahedron model of Budiansky and Kimmel (1987), letting the elastic behavior of a single regular dodecahedron (Fig. 1*a*), represent that of the aggregate of polyhedral cells in the foam. The faces of the dodecahedron are constrained to stay plane in the presence of P_c, and the walls have surface tension Γ only on one side. Under the uniform surface tractions that correspond to an arbitrary stress state σ_{ij}, the additional assumptions are made that the center deflection of each pentagonal face of the dodecahedron is the mean of the corner displacements (Fig. 1*b*); and that the displacements vary linearly within each of the five constituent congruent triangles of the pentagon. For small strains, the dodecahedron is isotropic, and two elastic constants suffice to relate the average strains ϵ_{ij} to σ_{ij}.

As stated by Stamenovic and Wilson, the bulk modulus K of the foam is dominated by the incompressibility of the trapped gas. The latter is of the order of the gas pressure P_i, which for terrestrial, room temperature foam is very many times larger than the capillary pressure P_c. Accordingly, since the shear modulus G is comparable to P_c, $G/K \ll 1$. So, in the calculation of G, we can let the trapped gas and the foam be incompressible, with Poisson's ratio $\nu = 1/2$.

We will now calculate the Young's modulus $E = \dot{\sigma}/\dot{\epsilon}$ by finding the averaged strain rate $\dot{\epsilon}$ in the direction of an external stress rate $\dot{\sigma}$ per unit area normal to σ (see Fig. 2). For the foam, the operative variational statement of equilibrium is

$$-P\delta V + \Gamma \delta S = \delta \pi \tag{4}$$

where π is the potential of the applied σ stresses. The subscripts on the capillary pressure P, the area S, and the volume V have been dropped, to indicate that they vary with the magnitude of σ. (We keep Γ constant, but, as will be shown, the results will be unchanged if Γ varies with S.) The changes in V and S depend only on four scalar displacements: w_1 and w_2 at B and C, respectively (see Fig. 2), in the direction of σ, and the magnitudes v_1, v_2 of the displacements at B, C normal to the σ direction. We have

$$V = V_o[1 + c_i y_i + \frac{1}{2} d_{ij} y_i y_j + 0(y^3)] \tag{5}$$

$$S = S_o[1 + c_i^* y_i + \frac{1}{2} d_{ij}^* y_i y_j + 0(y^3)] \tag{6}$$

[5]Professor. Division of Applied Sciences, Harvard University, Cambridge, MA 02138, Fellow, ASME.

[6]Senior Lecturer, Department of Agricultural Engineering, Technion, Haifa 32000, Israel.

Manuscript received and accepted by the ASME Applied Mechanics Division, Oct. 2, 1989.

[7]A reviewer of the present note informs us that this affine-deformation analysis was published earlier by Derjaguin (1933).

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

mechanically analogous to liquid foams, the change of parenchymal interfacial surface area with macroscopic shear strain should have a similar pattern as the change of foam surface area with the macroscopic shear strain.

Acknowledgment

The work was supported by National Heart, Lung, and Blood Institute, Grant HL 36958, and the Whitaker Foundation Grant.

References

Budiansky, B., and Kimmel, E., 1990, "On the Shear Modulus of Polyhedron-Cell Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp.

Kraynik, A. M., and Hansen, M. G., 1986, "Foam Emulsion Rheology: A Quasistatic Model for Large Deformations of Spatially-Periodic Cells," *Journal of Rheology*, Vol. 30, pp. 409–439.

Lai-Fook, S. J., 1979, "The Elastic Constants of Lung Parenchyma: The Effect of Pressure-Volume Hysteresis on the Behavior of Blood Vessels," *Journal of Biomechanics*, Vol. 12, pp. 757–764.

Princen, H. M., and Kiss, A. D., 1986, "Rheology of Foams and Highly Concentrated Emulsions. III. Static Shear Modulus," *Journal of Colloid and Interface Science*, Vol. 112, pp. 427–437.

Stamenović, D., and Wilson, T. A., 1984, "The Shear Modulus of Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 229–231.

On the Shear Modulus of Polyhedron-Cell Liquid Foam

B. Budiansky[5] and E. Kimmel[6]

Introduction

We consider liquid foam to be a cluster of closed cells bounded by thin, liquid walls, containing trapped gas. At a sufficiently high gas fraction, the cells are contiguous polyhedra of various sizes, shapes, and orientations (Aubert, 1986).

The pressure P_i of the trapped gas in each cell is slightly higher than the pressure P_e external to the foam. The overpressure, $P_c = P_i - P_e$, the so-called capillary pressure (Bikerman, 1973) is presumed to be balanced by surface tension forces in the cell walls.

For small external loading, in the absence of dissipation, the liquid foam may be idealized as an elastic isotropic material. Since the surface tension and the capillary pressure are constrained by equilibrium, the mechanical properties of the foam should depend only on the capillary pressure and the compressibility of the trapped gas.

By means of a Deer rheometer, Stamenovic and Wilson (1984) measured the elastic response of foam to an applied torque. Their measurements gave values of the ratio of the shear modulus G to capillary pressure P_c that ranged between .24 and .41, with an average value $G/P_c = .33$.

The somewhat larger result $G/P_c = .4$ was calculated theoretically by Stamenovic and Wilson on the basis of the assumption that the foam is a collection of randomly-oriented plates that deform in an affine manner.[7] However, recent studies of lung elasticity (Kimmel, Kamm, and Shapiro, 1987; Budiansky and Kimmel, 1987; Kimmel and Budiansky, 1989) suggest that the affine-deformation assumption is unduly constraining when applied to foam-like polyhedral assemblages. In this note we modify the analysis of Kimmel and Budiansky (1989) to derive the elastic moduli of foam.

Analysis

We first write the equilibrium connection between the capillary pressure P_c and the surface tension Γ per unit length that acts at each of the gas-liquid interfaces on each cell wall. By the principal of virtual work

$$P_c \Sigma \delta V_i = \Gamma \Sigma \delta S_i, \tag{1}$$

where the summations are over all the cells of the foam, V_i is the volume of the ith cell, and S_i is its boundary area. For self-similar (i.e., affine) virtual displacements, we have

$$\frac{\delta V_i}{V_i} = \frac{3}{2}\frac{\delta S_i}{S_i} \tag{2}$$

whence equation (1) implies

$$D \equiv \frac{\Gamma S_o}{P_c V_o} = \frac{3}{2} \tag{3}$$

where V_o and S_o are the average cell volume and boundary surface area. This relation has been derived repeatedly in various ways (Ross, 1969; Hoppin and Hildebrandt, 1979; Stamenovic and Wilson, 1984).

Next, we invoke the dodecahedron model of Budiansky and Kimmel (1987), letting the elastic behavior of a single regular dodecahedron (Fig. 1*a*), represent that of the aggregate of polyhedral cells in the foam. The faces of the dodecahedron are constrained to stay plane in the presence of P_c, and the walls have surface tension Γ only on one side. Under the uniform surface tractions that correspond to an arbitrary stress state σ_{ij}, the additional assumptions are made that the center deflection of each pentagonal face of the dodecahedron is the mean of the corner displacements (Fig. 1*b*); and that the displacements vary linearly within each of the five constituent congruent triangles of the pentagon. For small strains, the dodecahedron is isotropic, and two elastic constants suffice to relate the average strains ϵ_{ij} to σ_{ij}.

As stated by Stamenovic and Wilson, the bulk modulus K of the foam is dominated by the incompressibility of the trapped gas. The latter is of the order of the gas pressure P_i, which for terrestrial, room temperature foam is very many times larger than the capillary pressure P_c. Accordingly, since the shear modulus G is comparable to P_c, $G/K \ll 1$. So, in the calculation of G, we can let the trapped gas and the foam be incompressible, with Poisson's ratio $\nu = 1/2$.

We will now calculate the Young's modulus $E = \dot{\sigma}/\dot{\epsilon}$ by finding the averaged strain rate $\dot{\epsilon}$ in the direction of an external stress rate $\dot{\sigma}$ per unit area normal to σ (see Fig. 2). For the foam, the operative variational statement of equilibrium is

$$-P\delta V + \Gamma \delta S = \delta \pi \tag{4}$$

where π is the potential of the applied σ stresses. The subscripts on the capillary pressure P, the area S, and the volume V have been dropped, to indicate that they vary with the magnitude of σ. (We keep Γ constant, but, as will be shown, the results will be unchanged if Γ varies with S.) The changes in V and S depend only on four scalar displacements: w_1 and w_2 at B and C, respectively (see Fig. 2), in the direction of σ, and the magnitudes v_1, v_2 of the displacements at B, C normal to the σ direction. We have

$$V = V_o[1 + c_i y_i + \frac{1}{2} d_{ij} y_i y_j + 0(y^3)] \tag{5}$$

$$S = S_o[1 + c_i^* y_i + \frac{1}{2} d_{ij}^* y_i y_j + 0(y^3)] \tag{6}$$

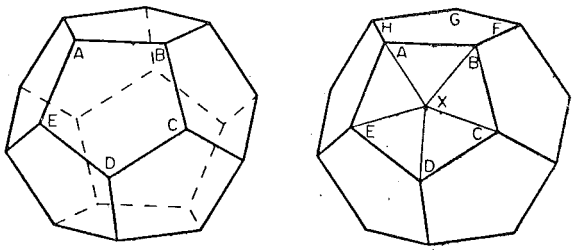
[5]Professor. Division of Applied Sciences, Harvard University, Cambridge, MA 02138, Fellow, ASME.

[6]Senior Lecturer, Department of Agricultural Engineering, Technion, Haifa 32000, Israel.

Manuscript received and accepted by the ASME Applied Mechanics Division, Oct. 2, 1989.

[7]A reviewer of the present note informs us that this affine-deformation analysis was published earlier by Derjaguin (1933).

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



(a) (b)

Fig. 1 (*a*) Regular dodecahedron, (*b*) subdivision of pentagonal faces

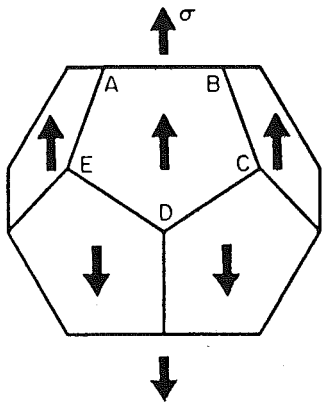


Fig. 2 Dodecahedron loaded by tension

$$\pi = \sigma V_o [f_i y_i]. \tag{7}$$

Here, $[y_1, y_2, y_3, y_4] = [w_1, v_1, w_2, v_2]/L_o$ where L_o is the length of each side of the dodecahedron at $\sigma = 0$, and S_o, V_0 are the bounding area and volume at $\sigma = 0$. The coefficients c_i, d_{ij}, and f_i were given by Budiansky and Kimmel (1987); $c_i^* = 2/3\ c_i$; and the d_{ij}^* were calculated numerically and tabulated by Kimmel and Budiansky (1989). Summation convention is used in equations (5)–(7).

At $\sigma = 0$, equation (4) reduces to the relation (3) connecting Γ, P_c, V_o, and S_o. Now differentiate (4) with respect to time, keeping σ_{ij} arbitrary, except for the additional incompressibility constraint $\delta V = 0$, and let $\sigma \to 0$. This leads to the five linear equations

$$-\left(\frac{\dot{P}}{P_c}\right) c_i + \left(\frac{3}{2} d_{ij}^* - d_{ij}\right) \dot{y}_j = \frac{\dot{\sigma}}{P_c} f_i \quad i = 1,2,3,4 \tag{8}$$

and

$$c_i \dot{y}_i = 0 \tag{9}$$

for the $\dot{y}_i$ and $\dot{P}$. Solving these equations, and evaluating $E = \dot{\sigma}/\dot{\epsilon}$ gives

$$E/P_c = .794. \tag{10}$$

With $\nu = 1/2$, we have $G = E/[2(1+\nu)] = E/3$, and so

$$G/P_c = .265. \tag{11}$$

Finally, we note that since pure shear produces no first-order change in area, G is not affected by variations of Γ with area of the form $\dot{\Gamma}/\Gamma = C\dot{S}/S$, where C is a constant. Accordingly, for $\nu = 1/2$, E is also independent of C.

Concluding Remarks

Our result for G/P_c is substantially lower than the value .4 calculated by Stamenovic and Wilson, but within their range of experimental measurements. Further tests are desirable.

Acknowledgment

This work was supported by the Office of Naval Research (Contract N00014-84-K-0510), the Division of Applied Sciences, Harvard University, and by the National Heart, Lung and Blood Institute of NIH (HL 33009).

References

Aubert, J. H., Kraynik, A. M., and Rand, P. B., 1986, "Aqueous Foams," *Scientific American*, Vol. 254, pp. 74–82.

Bikerman, J. J., 1973, *Foams*, Springer-Verlag, New York.

Budiansky, B., and Kimmel, E., 1987, "Elastic Moduli of Lungs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 351–358.

Derjaguin, B. V., 1933, "Die elastischen Eigenschaften der Schäume," *Kolloid-Z*, Vol. 64, pp. 1–6.

Hoppin, F. G., Jr., and Hildebrandt, J., 1979, "Mechanical Properties of the Lung," *Bioengineering Aspects of the Lung*, J. B. West, ed., Marcel Dekker, New York, pp. 83–162.

Kimmel, E., and Budiansky, B., 1990, "Surface Tension and the Dodecahedron Model for Lung Elasticity," *ASME Journal of Biomechanical Engineering*, Vol. 112, pp. 160–167.

Kimmel, E., Kamm, R. D., and Shapiro, A. H., 1987, "A Cellular Model of Lung Elasticity," *ASME Journal of Biomechanical Engineering*, Vol. 109, pp. 126–131.

Ross, S., 1969, "Bubbles and Foam—New General Law," *Industrial and Engineering Chemistry*, Vol. 61, No. 10, pp. 48–57.

Stamenovic, D., and Wilson, T. A., 1984, "The Shear Modulus of Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 229–231.

On the Crack Simulation and Solitons in the Lattice

N. F. Morozov[8] and M. W. Paucshto[8]

1 Introduction

As it is well known, discrete models are used for solving some problems in the continuum mechanics of solids.

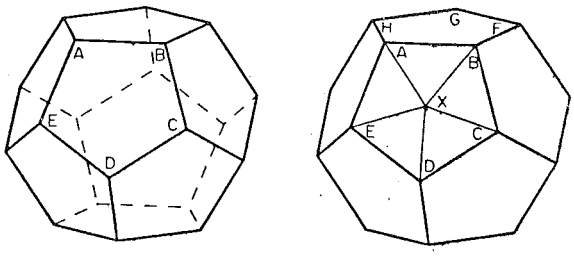
Newton (1686) used a discrete model for finding the optimal profile of a solid moving in gas, Cauchy (1829) developed discrete mechanics of solid consistently, and nowadays molecular dynamics methods are successfully employed in physics, biology, mechanics, and also in the theory of fracture. An example of calculations of discrete models of an elastic plane with a crack was performed by Ansorge et al. (1987). The calculations have demonstrated some quantitative and qualitative effects. The major disadvantage of the molecular dynamic technique is the restriction of the system to a finite number of structural elements. The basic question is, therefore, how large this number needs to be to simulate in a relevant way the properties of the corresponding infinite system (Dienes and Paskin, 1983).

In 1969 the theory of equilibrium cracks was developed by Novozilov. The author simulated the atomic forces and the zone of their action in a special way. Arising mathematical troubles are overcome by means of Leonov's and Panasuk's (1959) solution. Necessary conditions for the existence of equilibrium cracks have been formulated. Let $[l_-(p), l_+(p)]$ be the interval of equilibrium crack length corrresponding to any load p. It has been proved that $H = l_+(p)/l_-(p)$ is a material constant which does not depend on p.

[8]Department of Elasticity, Leningrad University, Leningrad, U.S.S.R.

Manuscript received by the ASME Applied Mechanics Division, Feb. 1989; final revision, Feb. 1990.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm



(a) (b)

Fig. 1 (*a*) Regular dodecahedron, (*b*) subdivision of pentagonal faces

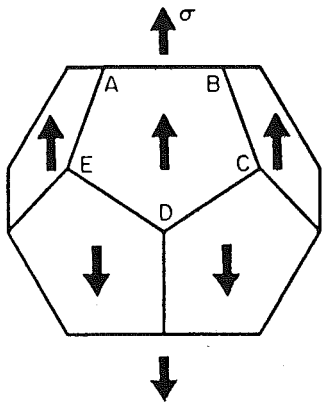


Fig. 2 Dodecahedron loaded by tension

$$\pi = \sigma V_o [f_i y_i]. \tag{7}$$

Here, $[y_1, y_2, y_3, y_4] = [w_1, v_1, w_2, v_2]/L_o$ where L_o is the length of each side of the dodecahedron at $\sigma = 0$, and S_o, V_0 are the bounding area and volume at $\sigma = 0$. The coefficients c_i, d_{ij}, and f_i were given by Budiansky and Kimmel (1987); $c_i^* = 2/3\ c_i$; and the d_{ij}^* were calculated numerically and tabulated by Kimmel and Budiansky (1989). Summation convention is used in equations (5)–(7).

At $\sigma = 0$, equation (4) reduces to the relation (3) connecting Γ, P_c, V_o, and S_o. Now differentiate (4) with respect to time, keeping σ_{ij} arbitrary, except for the additional incompressibility constraint $\delta V = 0$, and let $\sigma \to 0$. This leads to the five linear equations

$$-\left(\frac{\dot{P}}{P_c}\right) c_i + \left(\frac{3}{2} d_{ij}^* - d_{ij}\right) \dot{y}_j = \frac{\dot{\sigma}}{P_c} f_i \quad i = 1,2,3,4 \tag{8}$$

and

$$c_i \dot{y}_i = 0 \tag{9}$$

for the $\dot{y}_i$ and $\dot{P}$. Solving these equations, and evaluating $E = \dot{\sigma}/\dot{\epsilon}$ gives

$$E/P_c = .794. \tag{10}$$

With $\nu = 1/2$, we have $G = E/[2(1+\nu)] = E/3$, and so

$$G/P_c = .265. \tag{11}$$

Finally, we note that since pure shear produces no first-order change in area, G is not affected by variations of Γ with area of the form $\dot{\Gamma}/\Gamma = C\dot{S}/S$, where C is a constant. Accordingly, for $\nu = 1/2$, E is also independent of C.

Concluding Remarks

Our result for G/P_c is substantially lower than the value .4 calculated by Stamenovic and Wilson, but within their range of experimental measurements. Further tests are desirable.

Acknowledgment

This work was supported by the Office of Naval Research (Contract N00014-84-K-0510), the Division of Applied Sciences, Harvard University, and by the National Heart, Lung and Blood Institute of NIH (HL 33009).

References

Aubert, J. H., Kraynik, A. M., and Rand, P. B., 1986, "Aqueous Foams," *Scientific American*, Vol. 254, pp. 74–82.

Bikerman, J. J., 1973, *Foams*, Springer-Verlag, New York.

Budiansky, B., and Kimmel, E., 1987, "Elastic Moduli of Lungs," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 351–358.

Derjaguin, B. V., 1933, "Die elastischen Eigenschaften der Schäume," *Kolloid-Z*, Vol. 64, pp. 1–6.

Hoppin, F. G., Jr., and Hildebrandt, J., 1979, "Mechanical Properties of the Lung," *Bioengineering Aspects of the Lung*, J. B. West, ed., Marcel Dekker, New York, pp. 83–162.

Kimmel, E., and Budiansky, B., 1990, "Surface Tension and the Dodecahedron Model for Lung Elasticity," *ASME Journal of Biomechanical Engineering*, Vol. 112, pp. 160–167.

Kimmel, E., Kamm, R. D., and Shapiro, A. H., 1987, "A Cellular Model of Lung Elasticity," *ASME Journal of Biomechanical Engineering*, Vol. 109, pp. 126–131.

Ross, S., 1969, "Bubbles and Foam—New General Law," *Industrial and Engineering Chemistry*, Vol. 61, No. 10, pp. 48–57.

Stamenovic, D., and Wilson, T. A., 1984, "The Shear Modulus of Liquid Foam," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 229–231.

On the Crack Simulation and Solitons in the Lattice

N. F. Morozov[8] and M. W. Paucshto[8]

1 Introduction

As it is well known, discrete models are used for solving some problems in the continuum mechanics of solids.

Newton (1686) used a discrete model for finding the optimal profile of a solid moving in gas, Cauchy (1829) developed discrete mechanics of solid consistently, and nowadays molecular dynamics methods are successfully employed in physics, biology, mechanics, and also in the theory of fracture. An example of calculations of discrete models of an elastic plane with a crack was performed by Ansorge et al. (1987). The calculations have demonstrated some quantitative and qualitative effects. The major disadvantage of the molecular dynamic technique is the restriction of the system to a finite number of structural elements. The basic question is, therefore, how large this number needs to be to simulate in a relevant way the properties of the corresponding infinite system (Dienes and Paskin, 1983).

In 1969 the theory of equilibrium cracks was developed by Novozilov. The author simulated the atomic forces and the zone of their action in a special way. Arising mathematical troubles are overcome by means of Leonov's and Panasuk's (1959) solution. Necessary conditions for the existence of equilibrium cracks have been formulated. Let $[l_-(p), l_+(p)]$ be the interval of equilibrium crack length corrresponding to any load p. It has been proved that $H = l_+(p)/l_-(p)$ is a material constant which does not depend on p.

[8]Department of Elasticity, Leningrad University, Leningrad, U.S.S.R.

Manuscript received by the ASME Applied Mechanics Division, Feb. 1989; final revision, Feb. 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

A new model of brittle cracks was offered by Thomson et al. (1971, 1973). The potential energy of the model is obtained by summing up the potential energies of bending of elastic bars connected with springs and the straining energy of the springs. The springs are broken on reaching a critical length and uniwire restored when the corresponding shortened length is reached. The following strengths have been introduced: $p_+(n)$ is the maximal strength when n springs are broken while $p_-(n)$ is the minimal strength when n springs are broken. It has been proved that $R=p_+(n)/p_-(n)$ is the model's constant, which is independent of n.

Thomson's model seems more schematic than Novozilov's one. The credibility of Thomson's model is strengthened by calculations (Leora and Paukshto, 1987) and by its correspondence with Novozilov's model. It has been shown that $R=1/\sqrt{H}$ (Morozov and Paukshto, 1988).

The paper deals with the steady-state extension of a finite crack using Thomson's model.

2 Formulation of the Problem

Thomson's model allows us to analyze complicated problems of brittle fracture mathematically (Morozov and Paukshto, 1987). Let us consider the problem of a crack moving at speed v. The potential energy of this system is expressed by

$$U=\frac{1}{2}\{\gamma\Sigma(y_{j+1}-y_j)^2+\Sigma\Phi(y_j)\}, \tag{1}$$

where

$$\Phi(t)=\begin{cases}\beta r^2, & r\le u_0,\\ \beta u_0^2, & r>u_0.\end{cases} \tag{2}$$

Hence, one gets the following equations of motion

$$\gamma\{y_{j+2}-4y_{j+1}+6y_j-4y_{j-1}+y_{j-2}\}+\rho_j\beta y_j+m\ddot{y}_j=0, \tag{3}$$

where the function $\rho_j(t)=1$, if the j-spring is broken and $\rho_j(t)=0$, otherwise.

For the antiplane shear one has the analogous equations

$$m\ddot{y}_j-\gamma\{y_{j+1}-2y_j+y_{j-1}\}+\rho_j\beta y_j=0. \tag{4}$$

The function

$$\rho_j(t)=\begin{cases}0, & j-vt<0,\\ 1, & j-vt>0,\end{cases} \tag{5}$$

corresponds to a semi-infinite crack, moving with the velocity v. In this case the equation (4) is equivalent to Riemann's problem

$$\hat{y}_+(q)+\frac{4\sin^2(q/2)-v^2mq^2/\gamma}{4\sin^2(q/2)-v^2mq^2/\gamma+\beta/\gamma}\hat{y}_-(q)=0. \tag{6}$$

Note that the problem of a semi-infinite crack moving with a constant velocity v in the lattice (Slepian, 1984) is reduced to equation (6).

If $2k_0$ is a number of zeros and $2k_\beta$ is a number of poles of Riemann's problem factor, then k_0-k_β is a number of linearly independent solutions of the problem equation (6). In particular, if $k_\beta\ge k_0$, then there are no solutions, i.e., it is impossible for the crack to move with such velocity. A similar result is valid for the equation problem (3) with the function equation (5).

Before consideration of a soliton-type problem of a moving finite crack, we pay attention to the simple potentials such that the corresponding Newton's equations have soliton solutions. The first one is the potential of Frenkel-Kontorova's model:

$$\Phi(r)=\frac{\beta}{\pi}(1-\cos\pi r), \qquad \beta>0, \tag{7}$$

the second one is the potential of Toda's model:

$$\Phi(r)=\frac{\beta}{\alpha}(\exp(-\alpha r)-1)+\beta r, \qquad \alpha>0,\ \beta>0. \tag{8}$$

These potentials are the local elastic with nonbounded interaction ($\Phi'(r)\ne 0$ as $r->\infty$). Therefore, it is impossible to interpret the cracks by means of these potentials.

Let us take the function

$$\rho_j(t)=\begin{cases}1, & j-vt>l, \text{ or } j-vt<-l,\\ 0, & -l<j-vt<l.\end{cases} \tag{9}$$

Then the stationary solution of the system (3), i.e., the solution in the form

$$y(\tau)=y_j(t),\ \tau=j-vt, \tag{10}$$

satisfies the system of equations

$$y(\tau+2)-4y(\tau+1)+\left(6+v^2\frac{m}{\gamma}D_\tau^2\right)y(\tau)-4y(\tau-1)+y(\tau-2)+\rho(\tau)\frac{\beta}{\gamma}y(\tau)=0,\ \rho(\tau)=\rho_j(t). \tag{11}$$

Let $y(\tau)$ be a smooth enough function for $\tau\ne\pm l$ and such that

$$y(\pm l-0)=y(\pm l+0);\ Dy(\pm l-0)=Dy(\pm l+0). \tag{12}$$

Then applying Fourier's transformation to equations (11) we get

$$\left(16\sin^4\frac{w}{2}-v^2w^2m/\gamma+\beta/\gamma\right)\int_{-\infty}^{\infty}e^{iw\tau}y(\tau)d\tau=\frac{\beta}{\gamma}\int_{-\infty}^{\infty}(1-\rho(\tau))e^{iw\tau}y(\tau)d\tau. \tag{13}$$

3 Main Results

At this point, it is important to have the information about the zeros of the equation

$$16\sin^4(w/2)-v^2w^2m/\gamma+\beta/\gamma=0. \tag{14}$$

By applying the results of Pavlow (1973), we can prove the following statement:

Lemma. The asymptotic formula

$$w_k=k\pi\pm i\ln(|k|\pi v\sqrt{m/\gamma})+o(1),\ |k|\to\infty \tag{15}$$

is valid.

Moreover, the multiple roots of equation (14) are real. An algorithm to compute the roots was obtained by Sulimov and Paukshto (1990). Further, we shall consider those values of v and β/γ which only two real roots exist w_1, $-w_1$ ($w_1>0$).

Transform equality (13) to the form

$$y(\tau)=(Ay)(\tau)\equiv\frac{1}{2\pi}(\beta/\gamma)\int_{-\infty}^{\infty}e^{-i\tau w}K(w)\int_{-\infty}^{\infty}(1-\rho(\tau))e^{iw\tau}y(\tau)d\tau dw. \tag{16}$$

Here,

$$K(w)=1/(16\sin^4(w/2)-v^2w^2m/\gamma+\beta/\gamma).$$

It can be shown, that operator A is a compact operator in $L_2[-l, l]$. Its eigenvalues are condensed to $-\infty$ and there exists only a finite number of positive eigenvalues.

By means of symmetrization, $z(t)\equiv y(t)\sqrt{(1-\rho(t))}$, equation (16) is given by the form $z(t)=\sqrt{(1-\rho(t))}(A\sqrt{(1-\rho)}z(t))$. Now, the first eigenvalue can be found by means of Courant's principle

$$\lambda_1=\max\frac{a[y,y]}{\|y\|^2_{L_2[-1,1]}}, \tag{17}$$

where,

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$a[y, y] \equiv (Ay, y) = \frac{1}{2\pi}(\beta/\gamma)\int_{-\infty}^{\infty} K(t)\left|\int_{-l}^{l} e^{its} y(s)ds\right|^2 dt. \tag{18}$$

In the case when the function $K(t)$ is bounded and vanishes as $t \to \infty$, the asymptotic behavior of the eigenvalues of operator A in the space $L_2[-l, l]$ has been investigated by Widom (1963). In our case $K(t)$ has two simple poles, and the external integral (18) is the principal value of the integral.

The existence theorem for positive eigenvalues of form (18) is easily proved for $w_1 = k\pi/l$, where k is a positive integer. Really, if $y_0(t) = 1/\sqrt{(21)}$, as $t \epsilon [-l, l]$, then we have

$$a[y_0, y_0] = \frac{\beta}{2\pi\gamma}\int_{-\infty}^{\infty} 4K(w)\frac{\sin^2 1w}{21w^2}\,dw$$

$$= \int_{-w_1}^{w_1} \frac{\beta K(w)\sin^2 1w}{\pi\gamma 1w^2}\,dw$$

$$- \int_{|w|>w_1} \frac{\beta\sin^2 1w}{\pi 1\gamma w^2(v^2w^2m/\gamma - 16\sin^4(w/2) - \beta/\gamma)}\,dw. \tag{19}$$

We note that

$$w_1 = k\pi/1 \sim \sqrt{(\beta/\gamma)}\,(1/v)\,\sqrt{(\gamma/m)},\ \text{as}\ \beta/\gamma \to \infty. \tag{20}$$

The first integral on the right-hand side of equation (19) is positive and it's bounded from below by a constant independent of w_1. To estimate the second integral, we have

$$\int_{|w|>w_1} \frac{\beta\sin^2 lw\,dw}{\gamma\pi l(v^2w^2m/\gamma - 16\sin^4(w/2) - \beta/\gamma)}$$

$$\le \frac{\beta w^{-3}}{\gamma\pi l}\int_{|s|>1} \frac{\sin^2 sk\pi\,ds}{(s^2-1)s^2(v^2m/\gamma - 16g(s))} \tag{21}$$

$$\le O(1/w_1)(m/\gamma)(v^2/(\pi l))\int_{|s|>1} \frac{\sin sk\pi}{(s^2-1)}\frac{ds}{s^2},$$

$$g(s) = \frac{\sin^4(k\pi s/2l) - \sin^4(k\pi/2l)}{w_1^2(s^2-1)}.$$

The second integral will be order $1/|w_1|$ as $|w_1| \to \infty$.

Thus, we have proved the following: There exists a function $y_0(t)$ such that quadratic form $a[y_0, y_0]$ is positive, if β/γ is sufficiently large. Then, from equation (17), it follows that the first eigenvalue of this form is positive.

4 Conclusion

Finally, if $\lambda(v, \beta/\gamma, 1)$ is some eigenvalue of the quadratic form (18), then the equation $\lambda(v, \beta/\gamma, 1) = 1$ is the equation for finding the soliton-crack motion parameters. Results of this research can be applied to explain the effect of lowering the resistence of media by the motion of small inclusions (Chornij, 1987).

5 References

Ansorge, R., Grotkopf, U., Morozov, N., and Paukshto, M., 1987, "On the crack tip kinetics under impact loading," *Vestnik LGU*, Vol. 1, pp. 46–54.

Cauchy, A. L., 1829, "Sur les equations differentielles diequilibre," *Ex. de Math.*, Vol. 4.

Chornij, G., 1987, "Mechanism anomalno niskogo soprotivlenij pri dvisenii v tverdich sredach," *Dokl. AN USSR*, Vol. 292, No. 6, pp. 1324–1328.

Dienes, G. I., Paskin, A., 1983, "Computer Modeling of Crack," *Proc. of a NATO Advanced Research Institute on Atomistics of Fracture*, R. M. Latanision and J. R. Pickens, eds., Plenum Press, New York, pp. 671–704.

Leonov, M., and Panasjuk, V., 1959, "Razvitie melchaishikh trescin v tverdom tele," *Prikladnaj Matematika and Mechanika*, Vol. 5, No. 4.

Leora, S., and Paukshto, M., 1987," Analysis of fracture in 1-d model," *Probl. of Fracture Mech.*, V. G. Zubcaninov, ed., Kalinin, pp. 76–84.

Morozov, N., Paukshto, M., 1987, "Cracks dynamics discrete media," *Vestnik LGU*, Vol. 1, No. 3, pp. 67–71.

Morozov, N., and Paukshto, M., 1988, "On the lattice trapping," *Dokl. AN USSR*, Vol. 299, No. 2, pp. 323–325.

Newton, I., 1686, *Philosophical Naturalis Principia Mathematica*.

Novozilov, V., 1969a, "O neobkhodimom i dostatochnom uslovii khrupkoi procnosty," *Prikladnaj Matematika and Mechanika*, Vol. 33, No. 2, pp. 212–222.

Novozilov, V., 1969b, "K teorii ravnovesnikh trescin," *Prikl. Math. & Mech.*, Vol. 33, No. 5, pp. 797–812.

Pavlov, B., 1973, "Spectral analysis of diff. operator," *Problems of Math. Phis.*, M. S. Birman, ed., No. 6, pp. 101–119.

Slepian, L. L., 1984, "Antiplane problem of lattice crack," *Izvestia AN USSR, Mech.*, Tverdogo Tela, Vol. 258, No. 3, pp. 112–118.

Sulimov, M., Paukshto, M., 1990, "On the motion of semi-infinite crack," *Problems of Math. Phis.*, N. N., Uralceva, ed., to appear.

Thomson, R., Hsieh, C., and Rana, V., 1971, *J. Appl. Phys.*, Vol. 42, p. 3154.

Thomson, R., and Hsieh, C., 1973, *J. Appl. Phys.*, Vol. 44, p. 2051.

Widom, H., 1963, "Asymptotic behavior the eigenvalues of certen integral equation," *Trans. Amer. Math. Soc., 109*, Vol. 2, p. 278–295.

Bending of Plates With Rigid Inclusions Having Cuspidal Points

E. E. Gdoutos[9,10], M. A. Kattis[9], and M. Katsarou[9]

1 Statement of the Problem and Theoretical Background

Consider a plate of thickness h which is referred to a system of Cartesian coordinates $xy\delta$ with the x, y-plane coinciding with the midplane of the plate. The stress component σ_δ can be assumed to be zero throughout the plate. The stresses σ_x, σ_y, σ_{xy} and $\sigma_{x\delta}$, $\sigma_{y\delta}$ on an element located at (x, y, δ) are given by

$$\sigma_x = \frac{12\delta}{h^3}M_x \quad , \quad \sigma_y = \frac{12\delta}{h^3}M_y \quad , \quad \sigma_{xy} = \frac{12\delta}{h^3}M_{xy}$$

$$\sigma_{x\delta} = \frac{3(h^2-4\delta^2)}{2h^3}N_x \quad , \quad \sigma_{y\delta} = \frac{3(h^2-4\delta^2)}{2h^3}N_y \quad , \tag{1}$$

where M_x and M_y are the bending moments and M_{xy} is the twisting moment per unit length acting on any section of the plate parallel to the $x\delta$ and $y\delta$ planes, and N_x, N_y are the axial forces.

The plate is assumed to contain a perfectly bonded rigid inclusion of arbitrary shape which is mapped on the unit circle by the following equation

$$z = m(\zeta) = R\left(\zeta + \sum_{i=1}^{\rho}\frac{b_i}{\zeta^i}\right) \tag{2}$$

where $z(=x+iy)$ is the complex plane of the inclusion, $\zeta(=\exp(\xi+i\eta))$ is the mapping plane and R is a real and b_i are generally complex constants.

The moments M_ξ, M_η, $M_{\xi\eta}$ referred to the curvilinear coordinate system (ξ, η) and the displacement components u and v referred to the xy-system are expressed in terms of two complex potentials $W(\zeta)$ and $\mathrm{W}(\zeta)$ by the following equations (Savin, 1961)

$$M_\xi + M_\eta = -D(1+\nu)[W(\zeta) + \overline{W}(\bar{\zeta})] \tag{3a}$$

$$M_\xi + iM_{\xi\eta} = -\frac{D(1+\nu)}{2}[W(\zeta) + \overline{W}(\bar{\zeta})]$$

$$- \frac{D(1-\nu)}{2}\left[\frac{\bar{\zeta}m(\zeta)}{\zeta m'(\zeta)}\overline{W}'(\bar{\zeta}) + \frac{\bar{\zeta}\overline{m}'(\bar{\zeta})}{\zeta m'(\zeta)}\overline{\mathrm{W}}(\bar{\zeta})\right] \tag{3b}$$

[9]Department of Civil Engineering, Democritus University of Thrace, GR-671 00, Xanthi, Greece.

[10]Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, June 20, 1989; final revision, Jan. 25, 1990.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$a[y, y] \equiv (Ay, y) = \frac{1}{2\pi}(\beta/\gamma)\int_{-\infty}^{\infty} K(t)\left|\int_{-l}^{l} e^{its} y(s)ds\right|^2 dt. \tag{18}$$

In the case when the function $K(t)$ is bounded and vanishes as $t \to \infty$, the asymptotic behavior of the eigenvalues of operator A in the space $L_2[-l, l]$ has been investigated by Widom (1963). In our case $K(t)$ has two simple poles, and the external integral (18) is the principal value of the integral.

The existence theorem for positive eigenvalues of form (18) is easily proved for $w_1 = k\pi/l$, where k is a positive integer. Really, if $y_0(t) = 1/\sqrt{(21)}$, as $t \epsilon [-l, l]$, then we have

$$a[y_0, y_0] = \frac{\beta}{2\pi\gamma}\int_{-\infty}^{\infty} 4K(w)\frac{\sin^2 1w}{21w^2}\,dw$$

$$= \int_{-w_1}^{w_1} \frac{\beta K(w)\sin^2 1w}{\pi\gamma 1w^2}\,dw$$

$$- \int_{|w|>w_1} \frac{\beta\sin^2 1w}{\pi 1\gamma w^2(v^2w^2m/\gamma - 16\sin^4(w/2) - \beta/\gamma)}\,dw. \tag{19}$$

We note that

$$w_1 = k\pi/1 \sim \sqrt{(\beta/\gamma)}\,(1/v)\,\sqrt{(\gamma/m)},\ \text{as}\ \beta/\gamma \to \infty. \tag{20}$$

The first integral on the right-hand side of equation (19) is positive and it's bounded from below by a constant independent of w_1. To estimate the second integral, we have

$$\int_{|w|>w_1} \frac{\beta\sin^2 lw\,dw}{\gamma\pi l(v^2w^2m/\gamma - 16\sin^4(w/2) - \beta/\gamma)}$$

$$\le \frac{\beta w^{-3}}{\gamma\pi l}\int_{|s|>1} \frac{\sin^2 sk\pi\,ds}{(s^2-1)s^2(v^2m/\gamma - 16g(s))} \tag{21}$$

$$\le O(1/w_1)(m/\gamma)(v^2/(\pi l))\int_{|s|>1} \frac{\sin sk\pi}{(s^2-1)}\frac{ds}{s^2},$$

$$g(s) = \frac{\sin^4(k\pi s/2l) - \sin^4(k\pi/2l)}{w_1^2(s^2-1)}.$$

The second integral will be order $1/|w_1|$ as $|w_1| \to \infty$.

Thus, we have proved the following: There exists a function $y_0(t)$ such that quadratic form $a[y_0, y_0]$ is positive, if β/γ is sufficiently large. Then, from equation (17), it follows that the first eigenvalue of this form is positive.

4 Conclusion

Finally, if $\lambda(v, \beta/\gamma, 1)$ is some eigenvalue of the quadratic form (18), then the equation $\lambda(v, \beta/\gamma, 1) = 1$ is the equation for finding the soliton-crack motion parameters. Results of this research can be applied to explain the effect of lowering the resistence of media by the motion of small inclusions (Chornij, 1987).

5 References

Ansorge, R., Grotkopf, U., Morozov, N., and Paukshto, M., 1987, "On the crack tip kinetics under impact loading," *Vestnik LGU*, Vol. 1, pp. 46–54.

Cauchy, A. L., 1829, "Sur les equations differentielles diequilibre," *Ex. de Math.*, Vol. 4.

Chornij, G., 1987, "Mechanism anomalno niskogo soprotivlenij pri dvisenii v tverdich sredach," *Dokl. AN USSR*, Vol. 292, No. 6, pp. 1324–1328.

Dienes, G. I., Paskin, A., 1983, "Computer Modeling of Crack," *Proc. of a NATO Advanced Research Institute on Atomistics of Fracture*, R. M. Latanision and J. R. Pickens, eds., Plenum Press, New York, pp. 671–704.

Leonov, M., and Panasjuk, V., 1959, "Razvitie melchaishikh trescin v tverdom tele," *Prikladnaj Matematika and Mechanika*, Vol. 5, No. 4.

Leora, S., and Paukshto, M., 1987," Analysis of fracture in 1-d model," *Probl. of Fracture Mech.*, V. G. Zubcaninov, ed., Kalinin, pp. 76–84.

Morozov, N., Paukshto, M., 1987, "Cracks dynamics discrete media," *Vestnik LGU*, Vol. 1, No. 3, pp. 67–71.

Morozov, N., and Paukshto, M., 1988, "On the lattice trapping," *Dokl. AN USSR*, Vol. 299, No. 2, pp. 323–325.

Newton, I., 1686, *Philosophical Naturalis Principia Mathematica*.

Novozilov, V., 1969a, "O neobkhodimom i dostatochnom uslovii khrupkoi procnosty," *Prikladnaj Matematika and Mechanika*, Vol. 33, No. 2, pp. 212–222.

Novozilov, V., 1969b, "K teorii ravnovesnikh trescin," *Prikl. Math. & Mech.*, Vol. 33, No. 5, pp. 797–812.

Pavlov, B., 1973, "Spectral analysis of diff. operator," *Problems of Math. Phis.*, M. S. Birman, ed., No. 6, pp. 101–119.

Slepian, L. L., 1984, "Antiplane problem of lattice crack," *Izvestia AN USSR, Mech.*, Tverdogo Tela, Vol. 258, No. 3, pp. 112–118.

Sulimov, M., Paukshto, M., 1990, "On the motion of semi-infinite crack," *Problems of Math. Phis.*, N. N., Uralceva, ed., to appear.

Thomson, R., Hsieh, C., and Rana, V., 1971, *J. Appl. Phys.*, Vol. 42, p. 3154.

Thomson, R., and Hsieh, C., 1973, *J. Appl. Phys.*, Vol. 44, p. 2051.

Widom, H., 1963, "Asymptotic behavior the eigenvalues of certen integral equation," *Trans. Amer. Math. Soc., 109*, Vol. 2, p. 278–295.

Bending of Plates With Rigid Inclusions Having Cuspidal Points

E. E. Gdoutos[9,10], M. A. Kattis[9], and M. Katsarou[9]

1 Statement of the Problem and Theoretical Background

Consider a plate of thickness h which is referred to a system of Cartesian coordinates $xy\delta$ with the x, y-plane coinciding with the midplane of the plate. The stress component σ_δ can be assumed to be zero throughout the plate. The stresses σ_x, σ_y, σ_{xy} and $\sigma_{x\delta}$, $\sigma_{y\delta}$ on an element located at (x, y, δ) are given by

$$\sigma_x = \frac{12\delta}{h^3}M_x \quad , \quad \sigma_y = \frac{12\delta}{h^3}M_y \quad , \quad \sigma_{xy} = \frac{12\delta}{h^3}M_{xy}$$

$$\sigma_{x\delta} = \frac{3(h^2-4\delta^2)}{2h^3}N_x \quad , \quad \sigma_{y\delta} = \frac{3(h^2-4\delta^2)}{2h^3}N_y \quad , \tag{1}$$

where M_x and M_y are the bending moments and M_{xy} is the twisting moment per unit length acting on any section of the plate parallel to the $x\delta$ and $y\delta$ planes, and N_x, N_y are the axial forces.

The plate is assumed to contain a perfectly bonded rigid inclusion of arbitrary shape which is mapped on the unit circle by the following equation

$$z = m(\zeta) = R\left(\zeta + \sum_{i=1}^{\rho}\frac{b_i}{\zeta^i}\right) \tag{2}$$

where $z(=x+iy)$ is the complex plane of the inclusion, $\zeta(=\exp(\xi+i\eta))$ is the mapping plane and R is a real and b_i are generally complex constants.

The moments M_ξ, M_η, $M_{\xi\eta}$ referred to the curvilinear coordinate system (ξ, η) and the displacement components u and v referred to the xy-system are expressed in terms of two complex potentials $W(\zeta)$ and $\mathrm{W}(\zeta)$ by the following equations (Savin, 1961)

$$M_\xi + M_\eta = -D(1+\nu)[W(\zeta) + \overline{W}(\bar{\zeta})] \tag{3a}$$

$$M_\xi + iM_{\xi\eta} = -\frac{D(1+\nu)}{2}[W(\zeta) + \overline{W}(\bar{\zeta})]$$

$$- \frac{D(1-\nu)}{2}\left[\frac{\bar{\zeta}m(\zeta)}{\zeta m'(\zeta)}\overline{W}'(\bar{\zeta}) + \frac{\bar{\zeta}\overline{m}'(\bar{\zeta})}{\zeta m'(\zeta)}\overline{\mathrm{W}}(\bar{\zeta})\right] \tag{3b}$$

[9]Department of Civil Engineering, Democritus University of Thrace, GR-671 00, Xanthi, Greece.

[10]Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, June 20, 1989; final revision, Jan. 25, 1990.

Copyright © 1991 by ASME

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$\frac{\partial(u+iv)}{\partial\eta} = -\frac{\delta i}{2}[[W(\zeta)+\overline{W}(\bar{\zeta})]\zeta m'(\zeta) - [m(\zeta)\overline{W}'(\bar{\zeta})+\overline{m}'(\bar{\zeta})\overline{\mathrm{W}}(\bar{\zeta})]\bar{\zeta}], \tag{3c}$$

where D represents the flexural rigidity of the plate and is equal to $Eh^3/12(1-\nu^2)$ with E the modulus of elasticity of ν the Poisson's ratio. The following notation has been introduced:

$$\overline{F}(z)=\overline{F(\bar{z})}. \tag{4}$$

For the case when the plate is subjected to principal bending moments M_1 and M_2 at infinity, the functions $W(\zeta)$ and $\mathrm{W}(\zeta)$ take the form

$$W(\zeta) = -\frac{M_1+M_2}{2D(1+\nu)} + 0\left(\frac{1}{\zeta^2}\right) \tag{5a}$$

$$\mathrm{W}(\zeta) = \frac{M_2-M_1}{D(1-\nu)}e^{-2i\varphi} + 0\left(\frac{1}{\zeta^2}\right) \tag{5b}$$

where φ is the angle that the plane of the moment M_1 subtends with the x-axis.

2 Hilbert Formulation

The function $W(\zeta)$, which is defined in the infinite region L of the ζ-plane bounded by the unit circle, is now extended into the circular region $R(|\zeta|<1)$ of the ζ-plane by the equation

$$m'(\zeta)W(\zeta) = -m'(\zeta)\overline{W}\left(\frac{1}{\zeta}\right) + \frac{1}{\zeta^2}m(\zeta)\overline{W}'\left(\frac{1}{\zeta}\right) + \frac{1}{\zeta^2}\overline{m}'\left(\frac{1}{\zeta}\right)\overline{\mathrm{W}}\left(\frac{1}{\zeta}\right), \quad \zeta\in R. \tag{6}$$

By substituting equation (6) into equations (3b) and (3c), and taking on the limit of the resulting equations onto the unit circle, it is granted that

$$(M_\xi + iM_{\xi\eta})(\sigma) = -\frac{D(1+\nu)}{2}[W_0^L(\sigma)+\overline{W}_0^L(\bar{\sigma})] - \frac{D(1-\nu)}{2}[W_0^R(\sigma)+\overline{W}_0^L(\bar{\sigma})] \tag{7}$$

$$\frac{\partial(u+iv)(\sigma)}{\partial\eta} = -\frac{\delta i}{2}\sigma[W_0^L(\sigma)-W_0^R(\sigma)] \tag{8}$$

where it was put

$$W_0(\zeta)=m'(\zeta)W(\zeta). \tag{9}$$

In these equations $W^L(\sigma)$ and $W^R(\sigma)$ denote the limit values of $W(\zeta)$ as ζ tends to point σ of the unit circle from left and right, respectively, as one moves on the circle in the clockwise direction.

For situations of zero in-plane rotation of the inclusion, the boundary condition along the interface of the inclusion and the plate is expressed by

$$(u+iv)(\sigma)=\text{constant}. \tag{10}$$

Then, equation (10) gives

$$W_0^L(\sigma)-W_0^R(\sigma)=0. \tag{11}$$

Equation (11) constitutes the Hilbert problem. The unknown function has the form

$$W_0(\zeta)=A_0+\frac{A_1}{\zeta}+\cdots+\frac{A_{\rho+1}}{\zeta^{\rho+1}} \tag{12}$$

3 Determination of the Coefficients

Following the method used by Gdoutos and coworkers (1987, 1988, to appear) the $\rho+2$ unknown coefficients of equation (12) are determined by

$$A_0=-R\frac{M_1+M_2}{2D(1+\nu)}, \quad A_1=0 \tag{13}$$

$$\begin{aligned}
&\overline{A}_{\rho+1}-\rho K_\rho=0\\
&\overline{A}_\rho-(\rho-1)K_{\rho-1}=0\\
&\cdots\cdots\cdots\cdots\cdots\\
&\overline{A}_3-2K_2=0\\
&\overline{A}_2-K_1=B=R\frac{M_2-M_1}{D(1-\nu)}e^{-2i\varphi}
\end{aligned} \tag{14}$$

4 Local Stress Field Around Cuspidal Points

Consider now the cuspidal point $z_k=m(\zeta_k)$ along the inclusion boundary at which $m'(\zeta_k)=0$ and $m''(\zeta_k)\neq 0$. A Cartesian system x_1y_1 is attached at the point z_k in such a way that the x_1-axis is tangent to the two branches of the inclusion boundary at z_k.

Following the method used by Gdoutos and coworkers (1987, 1988, to appear) the moments M_ξ, M_η, $M_{\xi\eta}$ referred to the system $\xi\eta$ are expressed by

$$M_\xi = \frac{h^2K_1^{(k)}}{24\rho^{1/2}}\left(\frac{1-\nu}{1+\nu}\cos\frac{3\theta}{2}-\frac{5+3\nu}{1+\nu}\cos\frac{\theta}{2}\right) - \frac{h^2K_2^{(k)}}{24\rho^{1/2}}\left(\frac{1-\nu}{1+\nu}\sin\frac{3\theta}{2}-\frac{7+\nu}{1+\nu}\sin\frac{\theta}{2}\right), \tag{15a}$$

$$M_\eta = -\frac{h^2K_1^{(k)}}{24\rho^{1/2}}\left(\frac{1-\nu}{1+\nu}\cos\frac{3\theta}{2}+\frac{3+5\nu}{1+\nu}\cos\frac{\theta}{2}\right) + \frac{h^2K_2^{(k)}}{24\rho^{1/2}}\left(\frac{1-\nu}{1+\nu}\sin\frac{3\theta}{2}+\frac{7\nu+1}{1+\nu}\sin\frac{\theta}{2}\right), \tag{15b}$$

$$M_{\xi\eta} = \frac{h^2(1-\nu)K_1^{(k)}}{24(1+\nu)\rho^{1/2}}\left(\sin\frac{3\theta}{2}+\sin\frac{\theta}{2}\right) + \frac{h^2(1-\nu)K_2^{(k)}}{24(1+\nu)\rho^{1/2}}\left(\cos\frac{3\theta}{2}+3\cos\frac{\theta}{2}\right), \tag{15c}$$

where the complex stress intensity factor $K^{(k)}$ is defined by

$$K^{(k)}=K_1^{(k)}-iK_2^{(k)}=\frac{6D(1+\nu)}{h^2}\lim_{t\to 0}\sqrt{t}\,W(t) \tag{16}$$

and

$$\zeta=\zeta_k+h_k\sqrt{t}, \quad h_k=\frac{\sqrt{2}\zeta_k}{\sqrt{|m''(\zeta_k)|}} \tag{17}$$

with

$$t=\rho\exp(i\theta)\,(0<\rho<<R).$$

5 Application

The previously developed general equations will be applied to the particular case of a rectilinear inclusion of length $2l$.

For this case, the mapping function is

$$z=m(\zeta)=\frac{1}{2}\left(\zeta+\frac{1}{\zeta}\right) \tag{18}$$

and the complex potential $W(\zeta)$ takes the form

$$W(\zeta) = -\frac{M(s+1)}{2D(1+\nu)}\frac{\zeta^2+1}{\zeta^2-1}+\frac{M(s-1)}{D(1-\nu)}\frac{e^{2i\varphi}}{\zeta^2-1}, \tag{19}$$

where $M_1=M$ and $M_2=sM$.

The stress intensity factor $K^{(k)}$ at the cuspidal points of the inclusion, which coincide with its end points, is given by

$$K^{(I)}=-\frac{1}{4}\sigma_b\sqrt{21}\left[1+s-(s-1)\frac{1+\nu}{1-\nu}e^{2i\varphi}\right] \tag{20}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with

$$\sigma_b = \frac{6M}{h^2}. \tag{21}$$

Acknowledgment

The support of The General Secretariat of Research and Technology of The Greek Ministry of Industry, Energy and Technology is greatfully acknowledged.

References

Gdoutos, E. E., Eftaxiopoulos, D. A., and Kattis, M. A., 1987, "A Rigid Triangular Inclusion Partially Bonded to an Elastic Matrix," *Theoretical and Applied Fracture Mechanics*, Vol. 8, No. 3, pp. 173–185.

Gdoutos, E. E., Kattis, M. A., and Kourounis, C. G., 1988, "An Infinite Plate Weakened by a Curvilinear Hole," *Engineering Fracture Mechanics*, Vol. 31, No. 1, pp. 55–64.

Gdoutos, E. E., and Kattis, M. A., "A Rigid Curvilinear Inclusion Partially Bonded in an Elastic Matrix," *International Journal of Engineering Science*, to appear.

Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, Oxford, U.K.

On the Competition Between Failure and Instability in Progressively Softening Solids

A. Needleman[11]

Progressively softening solids admit the possibility of a complete loss of stress-carrying capacity and are also prone to shear localization instabilities. Quasi-static shearing of a planar block is used to illustrate, in a very simple context, circumstances where one or the other of these outcomes is favored. The block is characterized as a linear softening Kelvin-Voigt type solid and an imperfection in the form of a band of reduced stiffness is specified.

1 Introduction

The localization of deformation into narrow bands of intense straining is a characteristic feature of inelastic deformation. The physical mechanisms responsible for triggering localization include thermal softening, microfracturing processes, e.g., microcracking and microvoid nucleation and growth, and the textural effects that arise from the discreteness of slip systems in crystalline solids. In some, but not all, circumstances of interest, the mechanism that triggers localization gives rise to progressive softening.

The term progressive softening is used to refer to circumstances where, at least for certain deformation histories, the constitutive relation permits the stress to decrease monotonically to zero after the attainment of a maximum stress. For such a solid, failure in the sense of a complete loss of stress-carrying capacity is inevitable. It is also possible, however, for a shear localization instability to precede failure. The issue addressed here, in the context of a one-dimensional model problem, concerns the circumstances that favor one of these outcomes over the other. Essentially this same one-dimensional model has been used to address a variety of localization issues including necking of tensile bars (Hutchinson and Neale (1977)) and localized buckling in compressed structures (Tvergaard and Needleman (1980)) as well as adiabatic shear band development (Litonski (1977)) and Molinari and Clifton (1987)).

Proper modeling of progressive softening requires a framework that characterizes the actual softening process, for example through internal variables describing damage evolution or, for thermal softening, by consideration of the coupled thermomechanical problem. However, for the present purposes, a simple strain-softening constitutive relation suffices. As is now well appreciated, there are inherent difficulties associated with using classical (i.e., nonpolar) rate-independent constitutive relations to address localization issues (see, e.g., Needleman (1988)). Here, the material is characterized as a rate-dependent progressively softening solid.

2 Strain Localization in Simple Shear

A planar block, infinite in one direction and of unit width in the other, is subject to simple shear. The nonlinear kinematic effects that lead to possible normal strains or stresses are neglected, and attention is confined to quasi-static deformations. With x denoting the coordinate in the direction of unit width, equilibrium gives

$$\tau_{,x} = 0 \tag{1}$$

where $\tau(x, t)$ is the shear stress and a comma denotes partial differentiation with respect to the indicated variable.

The boundary conditions are

$$u(0,t) = 0 \qquad u(1,t) = \dot{\gamma} t \tag{2}$$

where $u(x, t)$ is the displacement and $\dot{\gamma}$ is a prescribed positive constant. In the circumstances considered here, the solution amplitude is proportional to $\dot{\gamma}$ and, for convenience, $\dot{\gamma}$ is set to unity.

The block is taken to be composed of two strips. One strip, termed the band, is in the interval $(1 - \beta) \le x \le 1$. The other strip, the material outside the band, is in the range $0 \le x < (1 - \beta)$. For what follows, the location of the strips and the number of strips is not of significance. All that matters is the total width of band material, β, and of material outside the band, $(1 - \beta)$.

Equilibrium, equation (1), requires that the stress state remain homogeneous so that

$$\tau_b = \tau_o \tag{3}$$

with $(\ )_b$ and $(\ )_o$ denoting quantities in the band and outside the band, respectively. Furthermore, since the solution is independent of x within each strip, quantities with subscripts $(\ )_b$ and $(\ )_o$ are functions of time only.

The material in each region is characterized by a constitutive relation of the Kelvin-Voigt form

$$\tau_b = \dot{\gamma}_b + b_b \gamma_b \qquad \tau_o = \dot{\gamma}_o + b_o \gamma_o \tag{4}$$

where $\gamma = u_{,x}$ and a superposed dot is an alternative notation for the partial derivative with respect to time. An identical positive viscosity has been specified inside and outside the band. The magnitude of the viscosity serves to set the time scale. If the coefficient of the strain rate term in the constitutive relation is denoted by $a_b = a_o = a$, then the role of a can be exhibited by replacing t with t/a.

The shear strain rate outside the band is $\dot{u}(\{1 - \beta\},t)/(1 - \beta)$, while that in the band is $[1 - \dot{u}(\{1 - \beta\},t)]/\beta$. Thus, the nominal shear strain rate is

$$1 = (1-\beta)\dot{\gamma}_o + \beta\dot{\gamma}_b. \tag{5}$$

Solving equation (5) for $\dot{\gamma}_o$ gives

$$\dot{\gamma}_o = \frac{1}{(1-\beta)} - \frac{\beta}{(1-\beta)}\dot{\gamma}_b. \tag{6}$$

Integrating equation (6) with respect to time gives an analogous expression for γ_o,

$$\gamma_o = \frac{t}{(1-\beta)} - \frac{\beta}{(1-\beta)}\gamma_b. \tag{7}$$

[11]Division of Engineering, Brown University, Providence, RI 02912. Fellow, ASME.

Manuscript received by the ASME Applied Mechanics Division, Sept. 5, 1989, final revision, Jan. 24, 1990.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

with

$$\sigma_b = \frac{6M}{h^2}. \tag{21}$$

Acknowledgment

The support of The General Secretariat of Research and Technology of The Greek Ministry of Industry, Energy and Technology is greatfully acknowledged.

References

Gdoutos, E. E., Eftaxiopoulos, D. A., and Kattis, M. A., 1987, "A Rigid Triangular Inclusion Partially Bonded to an Elastic Matrix," *Theoretical and Applied Fracture Mechanics*, Vol. 8, No. 3, pp. 173–185.

Gdoutos, E. E., Kattis, M. A., and Kourounis, C. G., 1988, "An Infinite Plate Weakened by a Curvilinear Hole," *Engineering Fracture Mechanics*, Vol. 31, No. 1, pp. 55–64.

Gdoutos, E. E., and Kattis, M. A., "A Rigid Curvilinear Inclusion Partially Bonded in an Elastic Matrix," *International Journal of Engineering Science*, to appear.

Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, Oxford, U.K.

On the Competition Between Failure and Instability in Progressively Softening Solids

A. Needleman[11]

Progressively softening solids admit the possibility of a complete loss of stress-carrying capacity and are also prone to shear localization instabilities. Quasi-static shearing of a planar block is used to illustrate, in a very simple context, circumstances where one or the other of these outcomes is favored. The block is characterized as a linear softening Kelvin-Voigt type solid and an imperfection in the form of a band of reduced stiffness is specified.

1 Introduction

The localization of deformation into narrow bands of intense straining is a characteristic feature of inelastic deformation. The physical mechanisms responsible for triggering localization include thermal softening, microfracturing processes, e.g., microcracking and microvoid nucleation and growth, and the textural effects that arise from the discreteness of slip systems in crystalline solids. In some, but not all, circumstances of interest, the mechanism that triggers localization gives rise to progressive softening.

The term progressive softening is used to refer to circumstances where, at least for certain deformation histories, the constitutive relation permits the stress to decrease monotonically to zero after the attainment of a maximum stress. For such a solid, failure in the sense of a complete loss of stress-carrying capacity is inevitable. It is also possible, however, for a shear localization instability to precede failure. The issue addressed here, in the context of a one-dimensional model problem, concerns the circumstances that favor one of these outcomes over the other. Essentially this same one-dimensional model has been used to address a variety of localization issues including necking of tensile bars (Hutchinson and Neale (1977)) and localized buckling in compressed structures (Tvergaard and Needleman (1980)) as well as adiabatic shear band development (Litonski (1977)) and Molinari and Clifton (1987)).

Proper modeling of progressive softening requires a framework that characterizes the actual softening process, for example through internal variables describing damage evolution or, for thermal softening, by consideration of the coupled thermomechanical problem. However, for the present purposes, a simple strain-softening constitutive relation suffices. As is now well appreciated, there are inherent difficulties associated with using classical (i.e., nonpolar) rate-independent constitutive relations to address localization issues (see, e.g., Needleman (1988)). Here, the material is characterized as a rate-dependent progressively softening solid.

2 Strain Localization in Simple Shear

A planar block, infinite in one direction and of unit width in the other, is subject to simple shear. The nonlinear kinematic effects that lead to possible normal strains or stresses are neglected, and attention is confined to quasi-static deformations. With x denoting the coordinate in the direction of unit width, equilibrium gives

$$\tau_{,x} = 0 \tag{1}$$

where $\tau(x, t)$ is the shear stress and a comma denotes partial differentiation with respect to the indicated variable.

The boundary conditions are

$$u(0,t) = 0 \qquad u(1,t) = \dot{\gamma} t \tag{2}$$

where $u(x, t)$ is the displacement and $\dot{\gamma}$ is a prescribed positive constant. In the circumstances considered here, the solution amplitude is proportional to $\dot{\gamma}$ and, for convenience, $\dot{\gamma}$ is set to unity.

The block is taken to be composed of two strips. One strip, termed the band, is in the interval $(1 - \beta) \le x \le 1$. The other strip, the material outside the band, is in the range $0 \le x < (1 - \beta)$. For what follows, the location of the strips and the number of strips is not of significance. All that matters is the total width of band material, β, and of material outside the band, $(1 - \beta)$.

Equilibrium, equation (1), requires that the stress state remain homogeneous so that

$$\tau_b = \tau_o \tag{3}$$

with $(\ )_b$ and $(\ )_o$ denoting quantities in the band and outside the band, respectively. Furthermore, since the solution is independent of x within each strip, quantities with subscripts $(\ )_b$ and $(\ )_o$ are functions of time only.

The material in each region is characterized by a constitutive relation of the Kelvin-Voigt form

$$\tau_b = \dot{\gamma}_b + b_b \gamma_b \qquad \tau_o = \dot{\gamma}_o + b_o \gamma_o \tag{4}$$

where $\gamma = u_{,x}$ and a superposed dot is an alternative notation for the partial derivative with respect to time. An identical positive viscosity has been specified inside and outside the band. The magnitude of the viscosity serves to set the time scale. If the coefficient of the strain rate term in the constitutive relation is denoted by $a_b = a_o = a$, then the role of a can be exhibited by replacing t with t/a.

The shear strain rate outside the band is $\dot{u}(\{1 - \beta\},t)/(1 - \beta)$, while that in the band is $[1 - \dot{u}(\{1 - \beta\},t)]/\beta$. Thus, the nominal shear strain rate is

$$1 = (1-\beta)\dot{\gamma}_o + \beta \dot{\gamma}_b. \tag{5}$$

Solving equation (5) for $\dot{\gamma}_o$ gives

$$\dot{\gamma}_o = \frac{1}{(1-\beta)} - \frac{\beta}{(1-\beta)} \dot{\gamma}_b. \tag{6}$$

Integrating equation (6) with respect to time gives an analogous expression for γ_o,

$$\gamma_o = \frac{t}{(1-\beta)} - \frac{\beta}{(1-\beta)} \gamma_b. \tag{7}$$

[11]Division of Engineering, Brown University, Providence, RI 02912. Fellow, ASME.

Manuscript received by the ASME Applied Mechanics Division, Sept. 5, 1989, final revision, Jan. 24, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Substituting equation (4) into equation (3) and using equations (6) and (7) gives the following equation for the evolution of γ_b

$$\frac{d\gamma_b}{dt} + \bar{b}\gamma_b = 1 + b_o t \tag{8}$$

where

$$\bar{b} = (1-\beta)b_b + \beta b_o \tag{9}$$

together with the initial condition $\gamma_o(0) = 0$.

For the special case of a homogeneous block with $\beta = 1$, the unique solution to equation (8), satisfying the initial condition, is $\gamma_b = t$. The uniform shear stress is $1 + b_b t$. A similar uniform solution pertains to $\beta = 0$, with $(\,)_b$ quantities replaced by $(\,)_o$ quantities. For the case with $\beta = 1$, the stress increases monotonically with time when $b_b > 0$, while for $b_b < 0$, the stress decreases monotonically with time and vanishes at $t = -1/b_b$.

In the present context, the unit stress at $t = 0$ can be regarded as having been attained through some unspecified history, with the constitutive relation equation (4) pertaining to the subsequent response. Also, for progressively softening materials, the loss of stress carrying capacity is regarded as the onset of failure and the end of the range of applicability of the constitutive relation. Hence, the failure criterion is $\tau_b = \tau_o = 0$. A localization criterion for rate-dependent solids is $\dot{\gamma}_b/\dot{\gamma}_o \to \infty$ (see, e.g., Hutchinson and Neale (1977)). By this criterion, a localization instability is associated with an increment of strain occurring in the band with no increase in strain outside the band. This can take place by $\dot{\gamma}_b$ becoming infinite or by $\dot{\gamma}_o$ vanishing with $\dot{\gamma}_b$ finite. The range of applicability of the constitutive relation equation (4) is also regarded as limited to the onset of a localization instability; equation (4) is not regarded as applicable in the post-localization regime.

In general, the solution to equation (8) is

$$\gamma_b(t) = t + (1-\beta)\left[\frac{(b_o - b_b)}{\bar{b}}\right]t - (1 - \beta)\left[\frac{(b_o - b_b)}{\bar{b}^2}\right][1 - \exp(-\bar{b}t)]. \tag{10}$$

Other field quantities of interest are

$$\dot{\gamma}_b(t) = 1 + (1-\beta)\left[\frac{(b_o - b_b)}{\bar{b}}\right] - (1 - \beta)\left[\frac{(b_o - b_b)}{\bar{b}}\right]\exp(-\bar{b}t) \tag{11}$$

$$\dot{\gamma}_o(t) = 1 - \beta\left[\frac{(b_o - b_b)}{\bar{b}}\right] + \beta\left[\frac{(b_o - b_b)}{\bar{b}}\right]\exp(-\bar{b}t) \tag{12}$$

$$\tau_b(t) = [1 + b_b t] + (1-\beta)\left[\frac{(b_o - b_b)}{\bar{b}}\right]b_b t + \beta(1-\beta)\left[\frac{(b_o - b_b)^2}{\bar{b}^2}\right][1 - \exp(-\bar{b}t)] \tag{13}$$

and from equation (3) $\tau_o = \tau_b$.

With $b_b \neq b_o$ there is an exponentially decaying component to the solution when $\bar{b} > 0$ and an exponentially growing component to the solution when $\bar{b} < 0$. In the subsequent discussion, the band material is considered to be less stiff than the material outside the band, i.e., $b_b < b_o$. However, both $b_o > 0$, a hardening material outside the band, and $b_o < 0$, a softening material outside the band, are considered.

In order to investigate the competition between failure, in the sense of loss of stress carrying capacity, and a localization instability, it is helpful to use equation (12) to rewrite equation (13) in the form

$$\tau_b(t) = [1 + b_b t]\left[1 + (1-\beta)\frac{(b_o - b_b)}{\bar{b}}\right] - \dot{\gamma}_o\left[(1-\beta)\frac{(b_o - b_b)}{\bar{b}}\right]. \tag{14}$$

Since $\dot{\gamma}_b$ remains finite for all finite t (as do all field quantities), $\dot{\gamma}_b/\dot{\gamma}_o \to \infty$ requires $\dot{\gamma}_o = 0$. From equation (12) this happens at $t = t_{\text{loc}}$, where

$$t_{\text{loc}} = \frac{-1}{\bar{b}}\ln\left[\frac{1}{\beta}\frac{b_b/b_o}{(b_b/b_o - 1)}\right]. \tag{15}$$

With $\bar{b}$, defined in equation (9), negative, t_{loc} is real and positive for

$$\frac{b_b/b_o}{b_b/b_o - 1} > \beta \tag{16}$$

while with $\bar{b}$ positive, real and positive t_{loc} requires

$$0 < \frac{b_b/b_o}{b_b/b_o - 1} < \beta. \tag{17}$$

The answer to the question of whether or not failure precedes localization hinges on the sign of $\tau_b(t_{\text{loc}})$. If $\tau_b(t_{\text{loc}}) > 0$, then localization precedes failure. On the other hand if $\tau_b(t_{\text{loc}}) < 0$, then failure has preceded the localization instability. Using equation (14), the dependence of $\tau_b(t_{\text{loc}})$ on b_b/b_o and β has the form

$$\tau_b(t_{\text{loc}}) = \frac{b_o^2}{\bar{b}^2}\left[\beta + (1-\beta)\frac{b_b}{b_o} - \frac{b_b}{b_o}\ln\left[\frac{1}{\beta}\frac{b_b/b_o}{(b_b/b_o - 1)}\right]\right] = \frac{b_o^2}{\bar{b}^2}F(b_b/b_o, \beta). \tag{18}$$

Recalling that $0 \le \beta \le 1$, consider the variation of $F(b_b/b_o, \beta)$ with b_b/b_o for fixed β. To make the discussion specific, first consider the situation where b_o is positive. For $b_b/b_o < -\beta/(1-\beta)$, $\bar{b}$ is negative and F is positive. For $-\beta/(1-\beta) < b_b/b_o < 0$ both $\bar{b}$ and F are positive (when $b_b/b_o = -\beta/(1-\beta)$, $\bar{b}$ and F vanish, but $\tau_b(t_{\text{loc}}) > 0$). Hence, for $b_b/b_o < 0$, localization precedes the loss of stress carrying capacity. In the range $0 < b_b/b_o < 1$, the hardening rate both in the band and outside the band is positive, but the band is less stiff than the surrounding material. Neither localization nor failure occur. The regime $b_b/b_o > 1$ corresponds to the band material being stiffer than the material outside the band. Next, consider $b_o < 0$ so that both the material inside the band and the material outside the band are softening. In this case a less stiff response for the band material requires $b_b < b_o$. Hence, the relevant interval is $b_b/b_o > 1$; $\bar{b}$ and F are negative and the loss of stress-carrying capacity precedes localization.

In summary, when the material outside the band hardens while the material inside the band softens, localization precedes failure; when the material inside the band and the material outside the band both soften, failure precedes localization. This feature is independent of the magnitude of the viscosity and, in particular, holds for an arbitrarily small positive viscosity.

Although $\dot{\gamma}_b/\dot{\gamma}_o \to \infty$ at $t = t_{\text{loc}}$, it is clear from equations (10) to (12) that a solution with $\dot{\gamma}_b/\dot{\gamma}_o < 0$ is available for $t > t_{\text{loc}}$. With $\bar{b} > 0$, $\dot{\gamma}_b/\dot{\gamma}_o \to b_b/b_b$ for large t, while with $\bar{b} < 0$, $\dot{\gamma}_b/\dot{\gamma}_o \to -(1-\beta)/\beta$ for large t. The significant negative straining required for $t > t_{\text{loc}}$ is often precluded by processes that are not even qualitatively modeled by equation (4), (e.g., elastic unloading-like phenomena).

Although whether localization or failure occurs first does not depend on band thickness, the separation between these events does, as can be seen in equation (18). Also, from equations (18) and (9), $\tau_b(t_{\text{loc}}) \to 0$ as $|b_b/b_o| \to \infty$. Hence, when the band material softens rapidly, localization and failure are nearly coincident, regardless of which occurs first.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The assumption of a linear viscous constitutive relation plays an important role in determining the response. Numerical experiments employing the viscoplastic constitutive relation used in Needleman (1988) indicate that power-law viscoplasticity gives rise to a strong preference for localization preceding failure, although in some cases, failure and localization take place nearly simultaneously. While it is expected that localization will generally precede failure, this model problem illustrates that such an outcome is not inevitable. In particular, when a linear viscous relation is an appropriate approximation, e.g., in certain creep regimes for a variety of solids, there is a rationale for the appearance of regions of progressively softening material within which localization does not precede failure, even though localization would be predicted based on a rate-independent analysis.

The present solution provides yet another illustration of the singular nature of the rate-independent limit. In the rate-independent limit the solution to the simple shearing problem is inherently nonunique, while for the Kelvin-Voigt type solid, the solution is unique. Furthermore, all field quantities remain finite for times up to and including the initiation of failure or the onset of a localization instability. Also, obtaining a numerical solution for the linear Kelvin-Voigt solid poses no special difficulties through the onset of localization and/or through the complete loss of stress carrying capacity. By way of contrast, numerical solutions for softening rate-independent solids exhibit an inherent mesh dependence (see, e.g., Needleman (1988)).

Acknowledgment

The support provided by the National Science Foundation (Solid Mechanics Program) through grant MSM-8618007 is gratefully acknowledged.

References

Hutchinson, J. W., and Neale, K. W., 1977, "Influence of Strain Rate Sensitivity on Necking under Uniaxial Tension," *Acta Metall.*, Vol. 25, pp. 839–846.

Litonski, J., 1977, "Plastic Flow of a Tube Under Adiabatic Torsion," *Bull. Acad. Polon. Sci., Sci. Tech.*, Vol. 25, pp. 7–14.

Molinari, A., and Clifton, R. J., 1987, "Analytical Characterization of Shear Localization in Thermoviscoplastic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 806–812.

Needleman, A., 1988, "Material Rate Dependence and Mesh Sensitivity in Localization Problems," *Comp. Meth. Appl. Mech. Eng.*, Vol. 67, pp. 69–85.

Tvergaard, V., and Needleman, A., 1980, "On the Localization of Buckling Patterns," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 613–619.

The Motion of a Sphere on a Rough Horizontal Plane

J.-Y. Yang[12]

1 Introduction

The study of the dynamic problem of a rotating rigid body on a plane is one that dates from D'Alembert's, Euler's, and Poisson's era (MacMillan, 1936; Routh, 1955). Since then, it has been examined along two lines: One has been about its stability, which was studied by Grammel (1952), Magnus (1971), Liu (1987), etc., and the other is on its analytic solution. J. A. Euler (Routh, 1955), for example, obtained the theorem that the direction and magnitude of the friction are constant. Slesser (MacMillan, 1936) gave the equations of the motion of a rigid body that is constrained to roll and pivot without sliding on a horizontal plane. Slesser's method is followed by Routh (Routh, 1955; MacMillan, 1936) in his discussion of the rolling and pivoting of a sphere on any surface. S. A. Chaplygin (1897, 1903) discussed the rolling of a rigid body and a nonhomogeneous sphere on a rough plane. There seems to be a mistake in J. A. Euler's deduction (MacMillan, 1936), which will be pointed out in this paper and the correct results for the motion of a sphere on a rough horizontal plane will be deduced.

2 The Differential Equations of the Motion of a Sphere on a Rough Horizontal Plane

On a rough horizontal plane, a sphere rotates about its axis, and it slides on this plane and nutates and proceeds. Let's suppose this sphere is exerted by its gravity and the reaction of the plane (containing the normal reaction and the friction).

In recent years, a coordinate system which doesn't rotate with the rotating top is used extensively for the symmetric top. Here, the same system is used. As Fig. 1 shows O is the center of the sphere, and $O\xi\eta\zeta$ is a system which translates with O. $Oxyz$ is a moving system which doesn't rotate with the rotating sphere, and the point of contact P is in the plane Oyz at all times. Or conversely, the plane Oyz follows the movement of P. Thus, the vector of angular velocity of the sphere ω is:

$$\omega = \dot{\varphi} + \dot{\theta} + \dot{\psi} = -\dot{\theta}\mathbf{i} - \dot{\varphi}\sin\theta\mathbf{j} + (\dot{\varphi}\cos\theta + \dot{\psi})\mathbf{k} \tag{1}$$

where φ, θ, and ψ are the angles of precession, nutation, and rotation, respectively. But the vector of angular velocity of the moving system $Oxyz$ Ω is as follows:

$$\Omega = \dot{\varphi} + \dot{\theta} = -\dot{\theta}\mathbf{i} - \dot{\varphi}\sin\theta\mathbf{j} + \dot{\varphi}\cos\theta\mathbf{k}. \tag{2}$$

If the angular velocity of rotation of the sphere is very large, then the change of nutation may be neglected, i.e., $\dot{\theta} = 0$, $\theta = \theta_o$. Hence, equations (1) and (2) will be as follows:

$$\omega = -\dot{\varphi}\sin\theta_o\mathbf{j} + (\dot{\varphi}\cos\theta_o + \dot{\psi})\mathbf{k} \tag{1'}$$

$$\Omega = -\dot{\varphi}\sin\theta_o\mathbf{j} + \dot{\varphi}\cos\theta_o\mathbf{k}. \tag{2'}$$

The radius vector of point P is as follows:

$$\mathbf{r}_p = \mathbf{OP} = \rho\sin\theta_o\mathbf{j} - \rho\cos\theta_o\mathbf{k}. \tag{3}$$

Hence, the velocity of point P is as follows:

$$\mathbf{V}_p = \mathbf{V}_o + \omega \times \mathbf{r}_p \tag{4}$$

where $\mathbf{V}_p$ and $\mathbf{V}_o$ are the velocity of point P and O in relation to a fixed plane, respectively. The projections of equation (4) in the system $Oxyz$ are as follows:

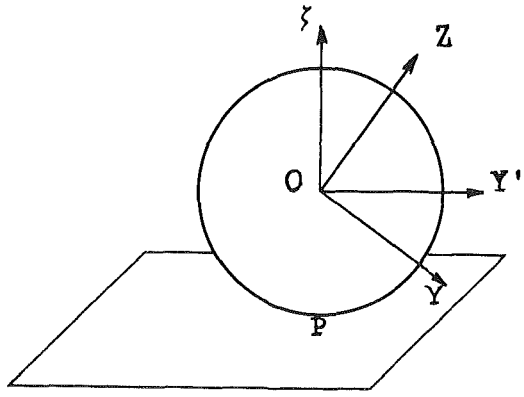


Fig. 1 The system of coordinates $O\xi\eta\zeta$ and $Oxyz$

[12]Department of Applied Physics, Shanghai Teachers' College of Technology, Shanghai, China.

Manuscript received by the ASME Applied Mechanics Division, Oct. 11, 1989; final revision, Jan. 17, 1990.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

The assumption of a linear viscous constitutive relation plays an important role in determining the response. Numerical experiments employing the viscoplastic constitutive relation used in Needleman (1988) indicate that power-law viscoplasticity gives rise to a strong preference for localization preceding failure, although in some cases, failure and localization take place nearly simultaneously. While it is expected that localization will generally precede failure, this model problem illustrates that such an outcome is not inevitable. In particular, when a linear viscous relation is an appropriate approximation, e.g., in certain creep regimes for a variety of solids, there is a rationale for the appearance of regions of progressively softening material within which localization does not precede failure, even though localization would be predicted based on a rate-independent analysis.

The present solution provides yet another illustration of the singular nature of the rate-independent limit. In the rate-independent limit the solution to the simple shearing problem is inherently nonunique, while for the Kelvin-Voigt type solid, the solution is unique. Furthermore, all field quantities remain finite for times up to and including the initiation of failure or the onset of a localization instability. Also, obtaining a numerical solution for the linear Kelvin-Voigt solid poses no special difficulties through the onset of localization and/or through the complete loss of stress carrying capacity. By way of contrast, numerical solutions for softening rate-independent solids exhibit an inherent mesh dependence (see, e.g., Needleman (1988)).

Acknowledgment

The support provided by the National Science Foundation (Solid Mechanics Program) through grant MSM-8618007 is gratefully acknowledged.

References

Hutchinson, J. W., and Neale, K. W., 1977, "Influence of Strain Rate Sensitivity on Necking under Uniaxial Tension," *Acta Metall.*, Vol. 25, pp. 839–846.

Litonski, J., 1977, "Plastic Flow of a Tube Under Adiabatic Torsion," *Bull. Acad. Polon. Sci., Sci. Tech.*, Vol. 25, pp. 7–14.

Molinari, A., and Clifton, R. J., 1987, "Analytical Characterization of Shear Localization in Thermoviscoplastic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 806–812.

Needleman, A., 1988, "Material Rate Dependence and Mesh Sensitivity in Localization Problems," *Comp. Meth. Appl. Mech. Eng.*, Vol. 67, pp. 69–85.

Tvergaard, V., and Needleman, A., 1980, "On the Localization of Buckling Patterns," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 613–619.

The Motion of a Sphere on a Rough Horizontal Plane

J.-Y. Yang[12]

1 Introduction

The study of the dynamic problem of a rotating rigid body on a plane is one that dates from D'Alembert's, Euler's, and Poisson's era (MacMillan, 1936; Routh, 1955). Since then, it has been examined along two lines: One has been about its stability, which was studied by Grammel (1952), Magnus (1971), Liu (1987), etc., and the other is on its analytic solution. J. A. Euler (Routh, 1955), for example, obtained the theorem that the direction and magnitude of the friction are constant. Slesser (MacMillan, 1936) gave the equations of the motion of a rigid body that is constrained to roll and pivot without sliding on a horizontal plane. Slesser's method is followed by Routh (Routh, 1955; MacMillan, 1936) in his discussion of the rolling and pivoting of a sphere on any surface. S. A. Chaplygin (1897, 1903) discussed the rolling of a rigid body and a nonhomogeneous sphere on a rough plane. There seems to be a mistake in J. A. Euler's deduction (MacMillan, 1936), which will be pointed out in this paper and the correct results for the motion of a sphere on a rough horizontal plane will be deduced.

2 The Differential Equations of the Motion of a Sphere on a Rough Horizontal Plane

On a rough horizontal plane, a sphere rotates about its axis, and it slides on this plane and nutates and proceeds. Let's suppose this sphere is exerted by its gravity and the reaction of the plane (containing the normal reaction and the friction).

In recent years, a coordinate system which doesn't rotate with the rotating top is used extensively for the symmetric top. Here, the same system is used. As Fig. 1 shows O is the center of the sphere, and $O\xi\eta\zeta$ is a system which translates with O. $Oxyz$ is a moving system which doesn't rotate with the rotating sphere, and the point of contact P is in the plane Oyz at all times. Or conversely, the plane Oyz follows the movement of P. Thus, the vector of angular velocity of the sphere ω is:

$$\omega = \dot{\varphi} + \dot{\theta} + \dot{\psi} = -\dot{\theta}\mathbf{i} - \dot{\varphi}\sin\theta\mathbf{j} + (\dot{\varphi}\cos\theta + \dot{\psi})\mathbf{k} \tag{1}$$

where φ, θ, and ψ are the angles of precession, nutation, and rotation, respectively. But the vector of angular velocity of the moving system $Oxyz$ Ω is as follows:

$$\Omega = \dot{\varphi} + \dot{\theta} = -\dot{\theta}\mathbf{i} - \dot{\varphi}\sin\theta\mathbf{j} + \dot{\varphi}\cos\theta\mathbf{k}. \tag{2}$$

If the angular velocity of rotation of the sphere is very large, then the change of nutation may be neglected, i.e., $\dot{\theta} = 0$, $\theta = \theta_o$. Hence, equations (1) and (2) will be as follows:

$$\omega = -\dot{\varphi}\sin\theta_o\mathbf{j} + (\dot{\varphi}\cos\theta_o + \dot{\psi})\mathbf{k} \tag{1'}$$

$$\Omega = -\dot{\varphi}\sin\theta_o\mathbf{j} + \dot{\varphi}\cos\theta_o\mathbf{k}. \tag{2'}$$

The radius vector of point P is as follows:

$$\mathbf{r}_p = \mathbf{OP} = \rho\sin\theta_o\mathbf{j} - \rho\cos\theta_o\mathbf{k}. \tag{3}$$

Hence, the velocity of point P is as follows:

$$\mathbf{V}_p = \mathbf{V}_o + \omega \times \mathbf{r}_p \tag{4}$$

where $\mathbf{V}_p$ and $\mathbf{V}_o$ are the velocity of point P and O in relation to a fixed plane, respectively. The projections of equation (4) in the system $Oxyz$ are as follows:

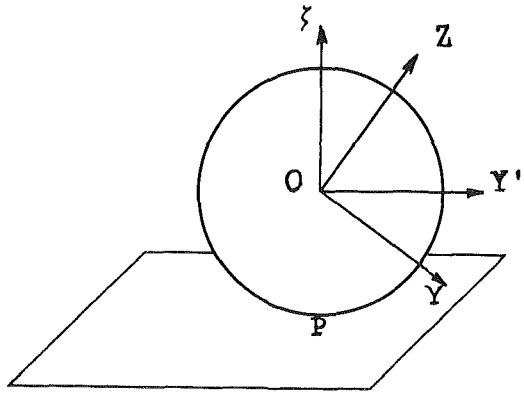


Fig. 1 The system of coordinates $O\xi\eta\zeta$ and $Oxyz$

[12]Department of Applied Physics, Shanghai Teachers' College of Technology, Shanghai, China.

Manuscript received by the ASME Applied Mechanics Division, Oct. 11, 1989; final revision, Jan. 17, 1990.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

$$V_{px} = V_{ox} - \rho\dot{\psi}\sin\theta_o \tag{5}$$

$$V_{py} = V_{oy} \tag{6}$$

$$V_{pz} = V_{oz}. \tag{7}$$

Making use of the principle of momentum,

$$Md/dt\mathbf{V}_o = \mathbf{W} + \mathbf{N} + \mathbf{F},$$

where $\mathbf{W}$ is the gravity, $\mathbf{N}$ the normal reaction, and $\mathbf{F}$ the friction. And, making use of the equation

$$d/dt\mathbf{V}_o = \tilde{d}/dt\mathbf{V}_o + \boldsymbol{\Omega}\times\mathbf{V}_o,$$

the following equations for the sphere are obtained:

$$M(\dot{V}_{ox} - V_{oz}\dot{\varphi}\sin\theta_o - V_{oy}\dot{\varphi}\cos\theta_o) = -\mu W V_{px}/V_p \tag{8}$$

$$M(\dot{V}_{oy} + V_{ox}\dot{\varphi}\cos\theta_o) = -\mu W V_{py}/V_p \tag{9}$$

$$M(\dot{V}_{oz} + V_{ox}\dot{\varphi}\sin\theta_o) = -\mu W V_{pz}/V_p. \tag{10}$$

In equations (8)–(10), owing to the fact that it is impossible for the mass center of the sphere to move in the vertical direction, N must equal W. Hence, in the right side of equations (8)–(10) N can be replaced by W, and

$$V_{p\zeta} = V_{o\zeta} = -V_{py}\sin\theta_o + V_{pz}\cos\theta_o = -V_{oy}\sin\theta_o + V_{oz}\cos\theta_o = 0. \tag{11}$$

Now, from the principle of moment of momentum about its center of mass other equations can be deduced. The moment of momentum of the sphere about point O. $\mathbf{G}_o$ is

$$\mathbf{G}_o = J\omega = -J\dot{\varphi}\sin\theta_o\mathbf{j} + J(\dot{\psi} + \dot{\varphi}\cos\theta_o)\mathbf{k} \tag{12}$$

where J is the moment of inertia of the sphere with respect to the diameter and equals $2M\rho^2/5$. Similarly,

$$d/dt\mathbf{G}_o = \tilde{d}/dt\ \mathbf{G}_o + \boldsymbol{\Omega}\times\mathbf{G}_o = \mathbf{OP}\times(\mathbf{F}+\mathbf{N}).$$

Other equations are deduced as follows:

$$-J\dot{\varphi}\dot{\psi}\sin\theta_o = -\mu W\rho(V_{pz}\sin\theta_o + V_{py}\cos\theta_o)/V_p \tag{13}$$

$$-J\ddot{\varphi}\sin\theta_o = \mu W\rho\cos\theta_o V_{px}/V_p \tag{14}$$

$$J(\ddot{\psi} + \ddot{\varphi}\cos\theta_o) = \mu W\rho\sin\theta_o V_{px}/V_p. \tag{15}$$

Before these equations are solved, let's first consider the component of velocity of the contact point P, $V_{py'}$ which is perpendicular to x-axis and ζ-axis and is in the Oyz-plane as shown in Fig. 1. Obviously,

$$V_{py'} = V_{oy'} = V_{pz}\sin\theta_o + V_{py}\cos\theta_o.$$

Making use of equation (11), $V_{py'}$ can be expressed as follows:

$$V_{py'} = V_{py}\sec\theta_o = V_{oy'} \tag{16}$$

and

$$V_p = \sqrt{V_{px}^2 + V_{py'}^2} = \sqrt{V_{px}^2 + V_{py}^2\sec^2\theta_o}. \tag{17}$$

Thus, equation (13) can be simplified

$$-J\dot{\varphi}\dot{\psi}\sin\theta_o = -\mu W\rho V_{py}\sec\theta_o/V_p. \tag{13'}$$

3 The Solution for the Angular Velocity of the Sphere

Making use of $[(15)\cos\theta_o - (14)\sin\theta_o]$, we obtain

$$J(\dot{\psi}\cos\theta_o + \dot{\varphi}) = \text{const or } \dot{\psi}\cos\theta_o + \dot{\varphi} = \dot{\psi}_o\cos\theta_o + \dot{\varphi}_o = C_1. \tag{18}$$

This equation shows that $G_{o\zeta}$, which is the component of $\mathbf{G}_o$ on the ζ-axis, is conservative. It is so because the external forces are parallel to the ζ-axis, or intersect it. Now, making use of $[(14)\cos\theta_o + (15)\sin\theta_o]$, we obtain

$$J\ddot{\psi}\sin\theta_o = \mu W\rho V_{px}/V_p. \tag{19}$$

As the result of $[(19)^2 + (13')^2]$, the following equation is obtained

$$J^2\ddot{\psi}^2\sin^2\theta_o + J^2\dot{\psi}^2\dot{\varphi}^2\sin^2\theta_o = \mu^2 W^2\rho^2. \tag{20}$$

Making use of equations (20) and (18), we can get the differential equation only containing ψ

$$J\ddot{\psi}\sin\theta_o = \pm\sqrt{(\mu W\rho)^2 - J^2\sin^2\theta_o(C_1 - \dot{\psi}\cos\theta_o)^2\dot{\psi}^2}. \tag{21}$$

Let's assume (Mathematical Handbook, 1979):

$$u = \sqrt{J\sin\theta_o\cos\theta_o}\,\dot{\psi} - C_1\sqrt{J\sin\theta_o/\cos\theta_o}/2 \tag{22}$$

$$a = \sqrt{\mu W\rho - (JC_1^2\sin\theta_o/4\cos\theta_o)},$$

$$b = \sqrt{\mu W\rho + (JC_1^2\sin\theta_o/4\cos\theta_o)}. \tag{23}$$

Then equation (21) has the following integral

$$\int_u^b du/\sqrt{(a^2+u^2)(b^2-u^2)} = \pm\int_t^{t_o}\sqrt{\cos\theta_o/J\sin\theta_o}\,dt. \tag{24}$$

However,

$$\int_u^b du/\sqrt{(a^2+u^2)(b^2-u^2)} = F(b/\sqrt{a^2+b^2},\ \arccos(u/b))/\sqrt{a^2+b^2} \tag{25}$$

becomes the standard elliptic integral of the first kind. Thus,

$$u = bcn(\sqrt{(a^2+b^2)\cos\theta_o/J\sin\theta_o}(t_o - t),\ b/\sqrt{(a^2+b^2)}) \tag{26}$$

whereby t_o it is meant that $u = b$ when $t = t_o$. It is known to all that $cn(z,k)$ is the elliptic cosine function. It is an even function. Thus, the mark "$\pm$" in front of z can be neglected. Let's return to $\dot{\psi}$

$$\dot{\psi} = C_1/2\cos\theta_o + bcn(\sqrt{(a^2+b^2)\cos\theta_o/J\sin\theta_o})(t_o - t),$$

$$b/\sqrt{a^2+b^2})/\sqrt{J\sin\theta_o\cos\theta_o}. \tag{27}$$

Making use of $\dot{\psi} = \dot{\psi}_o$ when $t = 0$, we can decide the unknown t_o from equation (27).

Since $cn(z,k)$ is the periodic function of variable z, $\dot{\psi}$ is the periodic function of variable t. From equation (18) we know $\dot{\varphi}$ is also a periodic function of t. And they have the same periods. The abovementioned results were not obtained by the predecessors (MacMillan, 1936; Routh, 1955).

It seems unusual that these solutions are periodic, because under the conditions of friction, it is generally believed that these solutions are decayed.

But, in experimental observations, $\dot{\psi}$'s and $\dot{\varphi}$'s periodicity during a considerably long period has really been discovered. Let's take, for example, a glass sphere which has a colored center rotates on a horizontal table. When we look down and askance at this sphere, we'll find its color periodically changes between those as shown in Fig. 2(a) and Fig. 2(b). Thus, the periodicity of the angular velocities of precession and rotation has been verified qualitatively by the experiment. However, in the experiment, a weak variance of the nutation can be observed. On the one hand, it shows our solutions still can not simulate the nature completely, on the other hand, it shows that neglecting the weak variance for the nutation is allowed and that it can reflect the periodicity of this phenomenon.

4 The Solution for the Velocity of Mass Center

Due to the constraint condition, in the equations of the velocity of mass center only two are actually independent. Thus, there is no harm in taking equations (8) and (9). Making use of equation (16), by transforming suitably, we can rewrite them in the following forms

$$\dot{V}_{ox} - V_{oy'}\dot{\varphi} = -\mu g V_{px}/V_p \tag{8'}$$

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

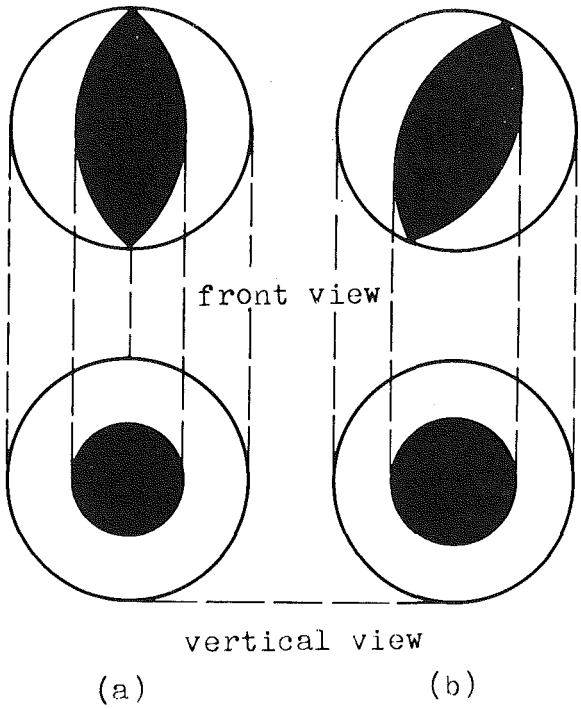


Fig. 2 The front and the vertical views

$$\dot{V}_{oy'} + V_{ox}\dot{\varphi} = -\mu g V_{py'}/V_p. \tag{9'}$$

As it is very hard to solve them, we are compelled to make a fresh start. Let's consider $\dot{V}_p$:

$$\dot{V}_p = (V_{px}\dot{V}_{px} + V_{py'}\dot{V}_{py'})/V_p.$$

Because $V_{px} = V_{ox} - \dot{\psi}\rho\sin\theta_o$, $V_{py'} = V_{oy'}$, the above equation can be transformed into:

$$\dot{V}_p = (V_{px}(\dot{V}_{ox} - \ddot{\psi}\rho\sin\theta_o) + V_{py'}\dot{V}_{oy'})/V_p.$$

Solving $\dot{V}_{ox}$, $\dot{V}_{oy'}$ and $\ddot{\psi}\sin\theta_o$ from equations (8′), (9′), and (19), and putting these into the above equation, we have

$$\dot{V}_p = -7\mu g/2. \tag{28}$$

This differential equation is unexpectedly simple. By integrating,

$$V_p = -7\mu g t/2 + V_{p_o}. \tag{29}$$

This result was obtained by J. A. Euler (MacMillan, 1936; Routh, 1955); however, his deduction seems to have a mistake, for the direction of $\dot{\mathbf{V}}_p$ doesn't parallel that of $\mathbf{V}_p$, otherwise the point of contact P of the sphere would always be on the plane. This mistake can also be found in equation (4) on p. 268 of MacMillan's work (MacMillan, 1936).

There is a time limit as V_p can't be negative. When $t > t_o' = 2V_{p_o}/7\mu g$, V_p should identically equal zero. At this time, the motion degrades to a inertial rotation or a pure roll. How do we solve the velocity of mass center when $t \le t_o'$? It is not required to solve equations (8′) and (9′), instead we can solve equations (13′) and (19). Now,

$$V_{oy'} = 2\dot{\varphi}\dot{\psi}\sin\theta_o\rho V_p/5\mu g \tag{30}$$

$$V_{ox} = \dot{\psi}\rho\sin\theta_o + 2\ddot{\psi}\sin\theta_o\rho V_p/5\mu g. \tag{31}$$

All of the quantities on the right side of equations (30) and (31) are now known. So far we have obtained all the solutions.

It is tenable only if $t \le t_o'$. These solutions should have the following qualities: $V_{oy'}$ decreases with the decrease of V_p, meanwhile it variates periodically. However, the first term of V_{ox} variates periodically, but the variance of the second is similar to that of $V_{oy'}$. Specifically, when V_p approaches zero, V_{ox} should have a periodical variance. It is also verified by the abovementioned experiments.

References

Chaplygin, S. A., 1987, *The Motion of a Rotating Rigid Body on a Horizontal Plane*, see *The Study on the Motion of Non-holonomic Systems* (1949), Moscow, pp. 9–27.

Chaplygin, S. A., 1903, *The Rolling of a Sphere on a Horizontal Plane*, see *The Study on the Motion of Non-holonomic Systems* (1949), Moscow, pp. 72–99.

Grammel, R., 1952, *Der Kreisel, seine Theorie und Seine Anwendungen*, Springer.

Liu, Y. Z., 1987, "The Stability of a Top Rotating on a Rough Horizontal Plane," *Acta Mechanica Sinica*, Vol. 19, No. 3, pp. 239–245.

MacMillan, W. D., 1936, *Dynamics of Rigid Bodies*, New York, pp. 257–276.

Magnus, E., 1971, *Kreisel Theorie und Anwendungen*, Springer.

Mathematical Handbook, 1979, Beijing, pp. 600–606, 640–648.

Routh, E. J., 1955, *The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies*, New York, pp. 165–186.

Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

Journal of Applied Mechanics

Book Reviews

Computational Mechanics of Probabilistic and Reliability Analysis, by W. K. Liu and T. B. Belytschko. Elmepress International, Lausanne, 1989. 622 pages. Price: $70.00.

REVIEWED BY P. SPANOS[1]

This book consists of a collection of 27 articles dealing with the subject of probabilistic engineering mechanics with special emphasis on the ensuing computational issues.

The book has two parts. The first part is meant as a tutorial on the basic aspects of probabilistic mechanics, whereas the second part is an overview of some of the applications that rely on probabilistic concepts for their treatment. The tutorial part starts with two expository chapters covering both basic and advanced concepts of structural reliability theory. Chapter 2, in particular, gives a clear and concise presentation of the first-order and second-order reliability methods, (FORM/SORM), as well as a thorough exposition of the various Monte Carlo simulation schemes that have proven to be useful in the context of structural reliability analysis. Chapters 3 and 4 present certain computational aspects associated with the analysis of structural systems with uncertain parameters. Chapter 3 deals with the dynamic analysis of such systems, whereas Chapter 4 treats them in the context of a perturbation-based stochastic finite element method. Chapter 5 is an attempt to couple the boundary element with a perturbation expansion for applying it to uncertain systems. In Chapter 6, a Neumann expansion is used to obtain statistics of the amplitude and frequency of uncertain vibratory systems. Useful truncation and convergence criteria associated with the expansion are also presented. The advanced mean value method for efficient estimation of the CDF is presented in Chapter 7. The NESSUS probabilistic finite element program is also reviewed in this chapter. The perturbation approach is again demonstrated in Chapter 8 through its application to the transient analysis of structures with uncertain damping. An improved perturbation approach is presented in Chapter 9. Here, an iterative perturbation scheme is reviewed that aims at improving the efficiency of the perturbation approach for the analysis of uncertain systems. Brief and sketchy presentations of directional simulation and importance sampling are presented in Chapter 10. The efficacy of these two methods in the reliability analysis of structural systems is indicated. Chapter 11 presents the coupling of the perturbation approach with the response surface approach for the reliability analysis of structural systems. The response surface approach is presented in Chapter 12 where it is applied to the analysis to uncertain nonlinear continua. A central composite design for the response surface is used for the computational experiments. In Chapter 13, the finite element implementation of the FORM and SORM methodologies is presented. The basic difficulties associated with such an implementation are pointed out, as well as various means of addressing them. A sketchy review of the representation of random fields is also presented. The first part of the book ends with a discussion of nonlinearity and randomness in mechanics based on the Adomian decomposition method.

The second part of the book presents a number of applications of probabilistic mechanics concepts in the context of computational mechanics. Six chapters, 15 through 20, address the problem of stochastic fracture mechanics and fatigue crack growth. State-of-the-art research in this important topic is presented. Random vibration analysis using Hermite moments is discussed in Chapter 21, and the first passage problem associated with stochastic structural dynamics is addressed in Chapter 22. Chapter 23 deals with hysteretic structural systems with special emphasis on the seismic reliability aspects of the problem. Chapters 24 and 25 feature more applications including the analysis of fluid-shell systems with random imperfections and the reliability analysis of the settlement of raft foundations. An algorithmic formulation is presented in Chapter 26 for the synthesis of linear two-phase uncertain systems. The Monte Carlo simulation method is used to simulate the system parameters. The last chapter of the book, Chapter 27, deals with the optimization of the performance reliability of manufacturing processes. A generalized reduced gradient method is implemented to maximize the system reliability.

In conclusion, the book attests to the timeliness of the rapidly evolving area of computational probabilistic mechanics. The choice of its material is somewhat misleading since it exclusively presents methods that are based on the FORM/SORM and the perturbation approaches for the analysis of uncertain structural systems. This choice leaves out a number of research contributions that are within the scope of the book. Still, this book is valuable to researchers and practitioners in the field of probabilistic mechanics.

Non-Destructive Testing, by B. Hull and V. John. Springer Verlag, New York, 1987. 144 pages.

REVIEWED BY TOSHIO MURA[2]

This is a thin (144 pages) introductory text which covers the underlying principles and some major applications of nondestructive inspection methods, liquid penetrant inspection, magnetic particle inspection, electrical testing, ultrasonic testing,

[1]L. B. Ryon Endowed Chair in Engineering, Department of Mechanical Engineering and Materials Science, Rice University, Houston, TX 77005.

[2]Department of Civil Engineering, Northwestern University, Evanston, IL 60208-3109.

Copyright © 1991 by ASME
Downloaded 03 May 2010 to 171.66.16.248. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm